Casualty Actuarial Society
E-Forum, Spring 2013
Volume 2

The CAS *E-Forum,* Spring 2013-Volume 2

The Spring 2013-Volume 2 edition of the CAS *E-Forum* is a cooperative effort between the CAS *E-Forum* Committee and various other CAS committees, task forces, or working parties.

Included in this volume is a research report on contingent capital, sponsored by the Committee on Valuation, Finance, and Investments. In addition, the CAS Committee on Reinsurance Research presents for discussion four papers prepared in response to the 2013 call for reinsurance papers. Some of the Reinsurance Call Papers will be discussed by the authors at the 2013 CAS Reinsurance Seminar on June 6-7, 2013, in Southampton, Bermuda.

Committee on Reinsurance Research

Mario E. DiCaro, *Chairperson*

Gary Blumsohn
Nebojsa Bojer
Michael A. Coca
Kirk Allen Conrad
Richard Jason Cook
Robert A. Giambo
Lilian Y. Giraldo
Evan W Glisson
Leigh Joseph Halliwell
Wei Juan Han
Robert L. Harnatkiewicz
James Anthony Heer
David L. Homer
Amanda Kisala
Richard Scott Krivo
Alex Krutov
Francois Lacroix
Apundeep Singh Lamba
Michael L. Laufer
Nicholas M. Leccese
Charles Letourneau
Li Ling Lin
Mundia Mubyana
Marc Lawrence Nerenberg
Raymond S. Nichols
James L. Norris
Gerard J. Palisi
Tracie L. Pencak
Anthony J. Pipia
John Bradley Raatz
Jaishan Rajendra
Nanxia Rao
Brett M. Shereck
Joel A. Vaag
Paul A. Vendetti
Ping Wang
Dominique Howard Yarnell
Ka Chun Yeung
Arthur J. Zaremba
Kan Zhong
Cheri Widowski, *Staff Liaison*

CAS *E-Forum,* Spring 2013-Volume 2

Table of Contents

E-Forum Committee

Windrie Wong, *Chairperson*

Cara Blank
Mei-Hsuan Chao
Mark A. Florenz
Karl Goring
Dennis L. Lange
Bryant Russell
Shayan Sen
Rial Simons
Elizabeth A. Smith, *Staff Liaison/Staff Editor*
John Sopkowicz
Zongli Sun
Betty-Jo Walke
Qing Wang
Yingjie Zhang

For information on submitting a paper to the *E-Forum*, visit http://www.casact.org/pubs/forum/.

Understanding Contingent Capital

Researcher:
Kailan Shang FSA, CFA, PRM, SCJP

February 2013

Prepared for:
Casualty Actuarial Society

Understanding Contingent Capital

Kailan Shang, FSA, CFA, PRM, SCJP

Abstract. This paper is a response to the Casualty Actuarial Society's request for proposals on "Contingent Capital." In light of the recent financial crisis, contingent capital, a type of hybrid security, is seen as an innovative way of recapitalization given the occurrence of a specified event, such as the capital adequacy ratio falling below the threshold. Although it has gained prominence among regulators, there are some doubts from market participants. The effectiveness of this automatic bail-in hybrid security is still too early to tell, given the limited market experience and unclearness of the impact on the share price when the conversion is triggered. The goal of this research is to explore the key features of contingent capital, its market, the appropriate pricing and valuation tools, and its application in insurance industry. It is hoped that the research will increase our understanding of contingent capital and facilitate the assessment of its value and risk.

Motivation. As a new alternative of raising capital automatically under stressed situations, contingent capital is expected to have more weight on insurers' balance sheets in the future. It is important for actuaries to understand contingent capital and have the necessary tools to assess its risk.

Method. This paper provides an overview of the contingent capital market, its features, and its potential impact. It also discusses the pricing and valuation models for certain contingent capital instruments. A case study is included to illustrate the quantitative analysis for a contingent capital instrument.

Results. A spreadsheet model is built and used in the case study. It is capable of pricing and valuing certain types of contingent capital. Quantitative risk analysis and model calibration function is also included. It could serve as good education materials to understand the role of contingent capital, quantify its risk, and assess its effectiveness of absorbing loss.

Conclusions. Contingent capital is a promising candidate to improve the capital position of the financial industry with a smaller cost than additional rights issuance. However, further analysis and testing are needed to find out the appropriate design and better understand its potential impact and related stakeholder behavior. There is still a journey to go before the success.

Availability. The spreadsheet "CONTINGENT CAPITAL QA TOOL" that illustrates the quantitative analysis of contingent capital is available, together with the report.

Keywords. Capital management, contingent capital, CoCo bond, systemic risk

Table of Contents

EXECUTIVE SUMMARY

The financial industry was quite successful in financial instruments innovations to meet the needs of investors. The wide development of the financial derivative market has been providing hedging tools to transfer financial risks for institutional investors. The insurance industry provides risk transfer as well, built on pooling the individual insurance risks together, and provides protection at a reasonable and affordable price. Using reinsurance and alternative risk transfer instruments that are normally linked with catastrophe risk, risk mitigation activities are prevalent in insurance industry.

However, they may not be sufficient to survive a financial crisis. Firms failed for different reasons but clearly they held insufficient capital for the risks they took. A huge amount of money was paid by taxpayers merely to keep some systematically important firms alive. The cost was supposed to be borne by the investors, however. The regulators and the financial industry are now taking steps to adopt a more stringent capital rule. Contingent capital caught a lot of attention as a candidate to strengthen the capital position, reduce the cost of financial crisis borne by the taxpayers, and limit the increasing cost of capital due to more stringent capital requirements. Generally speaking, contingent capital provides automatic recapitalization by converting the debt instrument to equity when the issuer is in trouble. The trigger of the conversion is based on a pre-specified condition, such as equity price dropping below a certain amount or regulatory capital ratio dropping below a certain level. This innovation links the automatic bail-in with the capital position of the company, covering a much broader scope of risks than ever before.

Several companies issued contingent capital instruments, either as a solution for the stressed financial condition or as an action to boost the capital buffer for future adverse events. There are many proposals of contingent capital design from the academic community, as well. The prevalence of different opinions of the appropriate trigger event for conversion implies the complexity of contingent capital and the immaturity of the market. There are concerns about the stakeholders' rational behaviors that may push the firm down to an even worse situation. This is exactly opposite to what the contingent capital is designed for. Other concerns include the softening of debt's disciplining power and its effectiveness. There is a lot to explore and test in the market before contingent capital becomes a widely accepted instrument for prudent capital management and risk management.

This paper introduces the background of contingent capital, its key features, its potential impact, and the models for pricing, valuation, and risk assessment. It explains the reasons for issuing contingent capital as well as the major concerns about it. It also includes a case study illustrating the pricing, valuation, and quantitative risk analysis of a sample contingent capital instrument.

1. INTRODUCTION

This paper is a response to the request for proposals on contingent capital by the Committee on Valuation, Finance, and Investments (VFIC) of the Casualty Actuarial Society (CAS).

1.1 Research Context

The insurance industry has been utilizing non-traditional capital instruments to transfer risks for a long time. Some of them help insurers absorb losses and retain their capital in adverse events. For example, catastrophe bonds or catastrophe equity put arrangements protect the insurers from catastrophe losses. Those instruments are normally related to insurance risk, such as natural disasters, mortality, and longevity.

In the recent financial crisis, systemic risk caught a lot of attention. Much discussion happened on how to prevent or mitigate systemic risk. Regulators are also changing their ways to regulate those too-big-to-fail, or systemically important financial institutions. Contingent capital, an innovative type of automatic bail-in hybrid security, is consideres a candidate for providing capital at a predetermined cost in stressed situations and for mitigating systemic risk. Contingent capital instruments are similar to non-traditional capital instruments used by insurance companies, except that the trigger is based on financial conditions instead of insurance risk. Different designs of contingent capital have been proposed and some of them are implemented. Although their effectiveness is still too early to tell, it is important for us to understand them and be equipped with knowledge and analytical tools for valuation and risk assessment.

1.2 Objective

The objective of this paper is to explore the key features and characteristics of contingent capital instruments, their effectiveness in risk transfer, and the pricing and valuation tools for them. A quantitative illustrative tool is made available for contingent capital evaluation and risk assessment. It is hoped that the tool will aid the actuarial community in understanding contingent capital from the perspective of risk transfer and capital management.

1.3 Outline

The remainder of the paper proceeds as follows. Section 2 gives an overview of the contingent capital market. Section 3 discusses the key features of contingent capital instruments. Section 4 presents their impact and effectiveness in risk mitigation and capital management. Section 5 explores modern finance theory and quantitative models used in pricing, valuation, and risk reward analysis. It is followed by a case study of evaluating a sample contingent convertible bond in Section 6. Section

7 concludes the paper.

2. CONTINGENT CAPITAL MARKET

Contingent capital instruments, also known as contingent convertible bonds (CoCo bonds), contingent surplus notes, or enhanced capital notes, provide a mechanism that automatically convert the instruments to equity upon the occurrence of a specified trigger event. These instruments began to attract attention and gain popularity during the 2008 financial crisis. Before that, insurance companies protected themselves from capital deficiency under stressed situations by reinsurance arrangements, hedging programs, and capital raising. Those seem effective when systemic risk is mild in the financial system. However, the recent financial crisis told us that when systemic risk is prevalent, the cost of raising capital may be unaffordable. Much higher liquidity risk and counterparty risk might still put the company in a weak solvency position. Contagion impact is material and the market and regulators have been looking for capital instruments that provide better insulation. Contingent capital appears to be the most promising solution, although doubts about it are not rare.

2.1 Market Overview

The insurance industry has been utilizing contingent capital instruments for around two decades. Catastrophe equity puts and contingent surplus notes are the most common types. Catastrophe equity puts[1] give the insurer the right to sell stocks at a fixed price in case a specified trigger event happens. Contingent surplus notes[2] give the insurer the right to issue surplus notes in exchange for liquid assets upon the occurrence of a predefined trigger event. The size of the transaction ranges from a few million dollars to around half a billion dollars. However, the trigger events, or, in other words, the risks from which the companies have been protected are normally catastrophe risk related. The term of the protection is also relatively short.

Contingent capital with a trigger event based on regulatory solvency ratio instead of insurance risk caught public attention in Lloyds Banking Group's exchange offer announced in November 2009. It intended to exchange certain existing securities[3] for enhanced capital notes or rights issue.[4]

---

[1] CEIOPS, "Insurance Linked Securities," 5.
[2] CEIOPS, "Insurance Linked Securities," 5.
[3] Existing securities subject to exchange offer comprised of "Upper Tier 2 securities in an aggregate principal amount of £2.52 billion, innovative Tier 1 securities in an aggregate principal amount of £7.68 billion and preference shares (or equivalents) with an aggregate liquidation preference of £4.09 billion", Lloyds Banking Group, "Exchange Offer," 2.

The enhanced capital notes (ECNs) will be converted to ordinary shares if the core tier 1 capital ratio falls below 5%. Not only the unprecedented size for contingent capital issuance but also the government's big stake in the company made this transaction very special. Contingent capital is considered more favorable than old-fashioned hybrid securities such as convertible bonds and preferred shares under stressed scenarios. These traditional hybrid securities are not as good as contingent capital instruments in loss absorbing.

In addition to conversion to equity if the trigger event happens, some other arrangements were also tried. In 2010, Rabobank issued a €1.25 billion 10-year Senior Contingent Note. Once the capital ratio falls below 7%, the face amount will be written down to 25% and paid back to investors. Liability value will be reduced if the trigger event happens, which effectively is a capital injection. This is different from contingent convertible bonds, where only the debt/equity ratio changes but the amount of capital remains the same if conversion is made at market price.

The €500 million deal of contingent convertible bonds between Allianz and Nippon Life in mid-2011 demonstrated the high interest of the insurance industry in using contingent capital to improve its capital position and reduce its risk exposure.

Not only banks and insurance companies but other financial institutes have used contingent capital. In the merger of Yorkshire Building Society and Chelsea Building Society in 2009, £200 million subordinated securities of Chelsea Building Society were planned in exchange for contingent convertible bonds. Once the core tier 1 capital ratio falls below 5%, they will automatically be converted into equity. [5]

2.2 Do We Need Contingent Capital?

Before diving into the details about contingent capitals, it is worth understanding the reasons for bringing contingent capital into the capital structure. From the regulators' perspective, it is hoped that contingent capital could solve the too-big-to-fail problem and reduce the loss paid by taxpayers instead of the investors. Compared to issuing new stocks, investors want to take advantage of the debt-like feature of the contingent capital: tax deductibility before the conversion and upfront and fixed recapitalization cost at conversion.

As seen in the financial crisis since 2008, many too-big-to-fail companies needed government

---

[4] The package was supposed to "(i) generate at least £7.5 billion in core tier one and/or nominal value of contingent core tier 1 capital through the Exchange Offer and/or related arrangements; and (ii) raise £13.5 billion (£13 billion net of expenses) by way of a Rights Issue, Lloyds Banking Group, "Exchange Offer," 1.

[5] A list of contingent capital instruments issued in the past few years is given in Goldman Sachs, "Contingent Capital Possibilities," 17-18.

bailout merely to survive. The bad outcomes of writing riskier business than what the available capital could support were borne by the taxpayers, not the investors who made the business decision. Higher capital requirement could certainly reduce the chance of default. However, as Bolton and Samama (2011) pointed out, the foreseen increased capital requirement for the banking industry according to Basel III may make it more difficult to earn the required return on equity (RoE) if the increased capital requirement needs to be met by the issuance of stocks. In addition, a sudden shift to a much more stringent capital requirement might also result in a credit crisis as the banks hold much less than the required capital as a buffer than previously required or have to raise capital to meet the capital requirement. Contingent capital seems to be a promising solution.

(1) As a debt instrument before conversion, it limits the increase in weighted average cost of capital (WACC). It will not cause the concern of higher required RoE in normal circumstances, which happens when financing with common equity.

(2) The tax deductibility of the debt instruments is also an argument for investors to utilize contingent capital in their financing. The disciplining power of creditors might also be preserved before conversion by maintaining the same leverage level as before.

(3) Firms normally try to sell troublesome assets and get rid of troublesome liabilities instead of issuing new stocks due to its high cost. Issuing contingent capital in good time fixes the recapitalization cost at a reasonable level in a future distressed situation. Apparently it is a cheaper way than raising capital in bad economic times.

(4) It can reduce the default probability without government bailout. Upon conversion, the capital base of the company will be increased so that it will have a stronger capital position than that before the conversion. The loss will be borne by the investors of contingent capital instead of the taxpayers. Therefore, it helps fulfill the goal of applying more stringent capital rules to too-big-to-fail firms. Bankruptcy and government bailout are very costly. Contingent capital can lower the chance of going through those expensive processes.

2.3 Designs of Contingent Capital

Despite the skepticism, there is a high and increasing interest in contingent capital instruments. Many designs of contingent capital have been proposed in trying to address the issues mentioned above. The key differences between possible designs so far concern the trigger event and the method of loss absorbency.

Scope of the trigger event. The trigger event can be based on the issuer's financial condition or on an industry-wide indicator. Industry-wide indicators, such as an aggregated market loss index or financial industry loss index, may be more suitable for mitigating systemic risk. However, it is hard to implement in an objective way and may convey an adverse message to the market.

Type of the trigger event. The trigger event can be related to the stock price, regulatory capital adequacy ratio such as core tier 1 ratio, credit condition such as credit default swap (CDS) spread, or even at the regulator's discretion. Double trigger contingent capital has also been proposed.

Level of the trigger point. Going-concern contingent capital normally has a high threshold, while gone-concern contingent capital has a low threshold, such as the point of non-viability. Besides the conversion of contingent capital, regulator's intervention is normally expected at the point of non-viability.

The method of loss absorbency. After the trigger event happens, contingent capital instruments will be converted to common equity, or have a write down of face amount and therefore liability. The impact is considered to be different. Write-down liability is similar to a capital injection, while conversion to equity is considered as capital restructure.

Details of those proposals and their different impacts will be explored in Section 3 and Section 4.

2.4 Stakeholder Analysis

Regulators have shown great interest in utilizing contingent capital to absorb losses under stressed conditions because it is expected to reduce the need of a government bailout and therefore taxpayers' support. After a sizable market is developed for contingent capital, the value of issuers, buyers, and existing stockholders and their roles in corporate governance will also be impacted. Rating agencies have also considered the rating methodology for this new type of hybrid securities and to what extent it can boost the issuer's financial strength.

2.4.1 Regulators

Regulators have been busy improving the capital adequacy rules to address the issues emerging from the financial crisis since 2008. In addition to a higher level of capital requirement, the qualification standard of hybrid securities to meet the additional capital requirement is also one of the major focuses. Although the future success of contingent capital is uncertain, it is very likely that contingent capital will become a part of the capital structure to meet regulators' requirements.

(1) In October 2010, Financial Stability Board (FSB) recommended that global systematically important financial institutions (SIFIs) should have higher loss absorbency.[6] One of the candidates that could be used to meet the stringent requirement is contingent capital which

[6] FSB, "Reducing the moral hazard," 3.

absorbs loss at the point of non-viability. It was later endorsed by the G20.

(2) Basel Committee on Banking Supervision (BCBS) released a rule about the additional loss absorbency requirement for global systematically important banks (GSIB) in late 2011. BCSC has considered different forms of contingent capital and decided the minimum requirements for going-concern contingent capital in order to meet additional requirements for GSIB. Contingent capital will have to be converted to Common Equity Tier 1 when the Common Equity Tier 1 falls below at least 7% of risk adjusted assets, which is a high threshold. It is also required to have a cap on the new shares and the full authorization of the issuers for an immediate conversion. The rule is expected to be phased in between 2016 and 2018 and become effective in 2019.

(3) The European Union amended its capital requirements directive (CRD) in 2009, known as CRD II, which highlighted the importance role of hybrid capital instruments in capital management. Instruments that absorb losses on a going-concern basis and that must be converted to core tier 1 capital are regarded as equity capital. In 2010, a consultation paper (CRD IV) that includes possible further changes was issued. It states that the European Commission will consider the potential need for all non-core tier 1 instruments to have a mandatory principal write-down or conversion feature, the potential triggers for conversion, and alternative mechanisms and triggers of contingent capital.[7]

(4) U.S. regulators are also interested in the idea of using contingent capital. As one of the provisions in the Dodd-Frank Wall Street Reform and Consumer Protection Act, Federal Reserve may establish heightened prudential standards for contingent capital requirement. It "*authorizes the Board to require each Board-supervised nonbank financial company and bank holding companies with total consolidated assets of $50 billion or more to maintain a minimum amount of contingent capital convertible to equity in times of financial stress.*"[8] The Fed is in discussions with bankers. Unlike the existing contingent capital deals that have trigger events related to the issuer's own capital ratio, the Fed is also exploring a system wide trigger.

(5) Office of the Superintendent of Financial Institutions Canada (OSFI) issued its final advisory on Non-Viability Contingent Capital. Seen as a fast movement on implementation of Basel III, it requires that the regulatory capital of all federally regulated deposit-taking institutions (DTIs) must have loss absorbing quality when the DTI fails. All the non-common Tier 1 and

[7] COMMISSION SERVICES STAFF WORKING DOCUMENT, "POSSIBLE FURTHER CHANGES," 18-19&23.
[8] Sec 165, "Enhanced supervision," H.R.4173

Tier 2 capital must satisfy the requirement for non-viability contingent capital (NVCC). Unlike CoCo bonds, the trigger event of NVCC is dependent on the regulators' announcement, either from OSFI, or federal/provincial government. Though it sounds like regulators have a lot of discretion, non-viability is determined based on clearly defined criteria which would mitigate the chance of discretion. Interestingly, this is more like a low trigger level, gone concern type of contingent capital compared to the high trigger level, going-concern contingent capital required for GSIB by BCBS.

(6) Under the Solvency II framework, contingent capital with appropriate feature can be classified as ancillary own fund[9] (AOF). AOF can be used to meet the solvency capital requirement (SCR) but not the minimum capital requirement (MCR). However, according to the directive, the total amounts and the amount for each AOF item are subject to supervisory approval. The recoverability, legal form, and any past exercise need to be taken into account when determining the amount qualified for AOF.[10]

(7) In August 2010, National Association of Insurance Commissioners (NAIC) Securities Valuation Office (SVO) reported on contingent capital securities. Considering that there is no agreement on the design of the trigger event, the task force did not draw any conclusion but decided to continue monitoring the development of contingent capital.

It is clear that contingent capital is one of the priorities of regulators regarding regulating SIFI but there is still work that needs to be done for further assessment and refinement. Regulators under different jurisdictions may also have different opinions regarding the details.

2.4.2 Issuers

In order to meet more stringent capital requirements, financial institutions can either raise more common equity or issue contingent capital instruments that must convert to common equity upon the occurrence of a trigger event.

Despite the greater complexity and the higher uncertainty of contingent capital, its potential cost is lower than that of common equity. This could increase the capacity for loss absorption and attract more issuers. In addition, the conversion of contingent capital to common equity normally means a dilution of existing shareholders' value. It will discourage shareholders from taking excessive risks above its capacity in the fear of conversion. There is also a chance that contingent capital will be made mandatory by regulators.

---

[9] "Ancillary own funds are items of capital other than basic own funds which can be called up to absorb losses." CEIOPS, "QIS5," 308.
[10] "Directive 2009/138/EC," 48.

2.4.3 Debt holder

The investors of contingent capital may be either holders of hybrid securities which will be exchanged for contingent capital as part of the company's restructuring plan or buyers of newly issued contingent capital instruments. In the Lloyds Banking Group's exchange offer in 2009, higher returns and immediate coupon payments motivated the exchange. New buyers may seek for high return and the potential gain from recovery after the conversion of contingent capital.

Like other debt holders, contingent capital investors would discipline the risk taking of stockholders. It is true especially when the conversion price is high, or the chance of recovery is low. However, on the other hand, when the company's financial condition is near the trigger event, in order to get a lower conversion price for conversion at par, investors may short the stocks to drag the stock price further down.

2.4.4 Shareholders

The shareholders' value and role will change if contingent capital becomes an important component of the financial institutions' capital. In the long run, effective contingent capital can improve corporate governance, partly solve agency problem, limit excessive risk taking, and reduce the cost of capital. However, during the conversion of contingent capital, the existing shareholders' value is often diluted. This may encourage more prudent risk taking activities. But the expectation of conversion in the near future will lead to more stock selling, which further lowers the stock price. This downward spiral will exacerbate the financial condition.

2.4.5 Rating Agencies

Rating agencies updated their rating methodologies of hybrid securities in the light of expanding contingent capital markets. It is a critical factor to consider when setting the price of contingent capital instruments. The view of rating agencies is also important for financial strength ratings when contingent capital becomes sizable in loss absorbency.

Debt Instrument

Contingent capital instruments that are rated as debt instruments normally receive ratings lower than investment grade.

In 2009, S&P issued a rating criterion for contingent capital. Contingent capital is defined as "*debt and hybrid securities that contain triggers that convert them into equity or some other Tier-1 instrument.*"[11] S&P believes that the proposed contingent capital increases the risk of loss to the investors, compared to plain vanilla bonds. Contingent capital securities would receive lower credit ratings than similar ones

[11] "Standard & Poor's Ratings Services Criteria Regarding Contingent Capital," 2.

without the conversion option. Conversion of capital securities will be treated as default.

Moody's classifies contingent capital into 3 groups: rate, no rate, and may rate. Moody's only rates "*securities that feature triggers for conversion that are credit-linked, objective and measurable and where the impact of conversion can be estimated.*" For triggers that are at the issuer's discretion or unrelated to the financial condition of the issuers, there will be no rating. When assigning the rating, the major considerations are "*the risk that the 'host' security might absorb losses for a 'going' concern, and the expected loss severity upon conversion based on the conversion ratio and the likely value that would be received by investors at that point in time. Important considerations would include the type and transparency of the trigger, how it is calculated, and over what time horizon.*" Moody's is also concerned with contingent capital instruments that have trigger events based on a regulatory capital ratio. "*Because many banks currently operate in rapidly changing regulatory and political environments, a lack of clarity on legal triggers and an overall resolution framework would prevent Moody's from assigning a rating at this time.*"[12]

<u>Equity Instrument</u>

When doing financial analysis, S&P considers a high trigger level and timely conversion critical for qualifying for an equity instrument, as stated below.

"*The conversion would need to happen early enough in the issuer's credit deterioration to be able to make a difference to that decline. A trigger level set at the regulatory capital threshold--or very close to it--is generally insufficient in our view to warrant equity-like treatment in advance of actual conversion. Similarly, if the conversion mechanisms allow for a potential significant lag after the trigger breach, we would not view the security as equity-like. Such lags could arise from stipulated delays or from pragmatic considerations, such as infrequent trigger measurement dates.*"[10]

Moody's determines the amount of equity credit for contingent capital based on their structures. Moody's thinks that "*triggers have generally not proven to be fail-safe in terms of their ability to accurately identify credit deterioration.*"[11] It is also difficult to ignore its similarity to debt instruments such as the fixed coupon rate and the need of refinance after the maturity of contingent capital.

Rating agencies focus on whether the trigger event is clear and objective, and whether it will result in timely conversion, which is the key to determine its effectiveness of loss absorption.

2.5 Some Doubts about Contingent Capital

Most of the contingent capital securities issued to date have trigger events linked to the capital ratio of the issuer. Given that the major goal of regulators and issuers is to reduce systemic risk exposure, there are doubts about the effectiveness of risk mitigation and the future of contingent

---

[12] Moody's, "Rating Considerations," 1-6.

capital. Investors and researchers have voiced many suspicions about the contingent capital instruments.

(1) Most of the trigger events used so far are based on the company's own core tier 1 capital ratio, a regulatory capital adequacy measure. This might not provide timely conversion considering the fast downward slide during a crisis.[13] There are also concerns about using the capital ratio, a measure based on market value. In a stressed situation, the owners of contingent capital, anticipating a drastic drop in equity price and occurrence of the trigger event in the near future, may short the stocks and exert pressure on the stock price. By doing this, they can get a lower conversion price. Stockholders may sell the stocks with the expectation of such behavior. This downward spiral may totally devour the benefits of the conversion and dilute the value of existing stockholders.

(2) To mitigate systemic risks and reduce the need of government bailouts for systematically important financial institutions, the market size of contingent capital needs to be big enough.[14] New features of contingent capital cause difficulty and uncertainty for both the pricing and valuation of this new type of hybrid security. Its higher risk and the lack of knowledge and experience may daunt many investors. There are also investors who have an investment policy that disallows equity market investment or have a limit on equity allocation. They may not be able to invest in contingent capital. In addition, when contingent capital is converted to common equity, they may be forced to sell the stocks, which may have a big market impact and more loss. This is also a potential impediment for the development of the contingent capital market.

(3) There have been hot discussions about contingent capital's impact on systemic risk. Some argue that the trigger event should be based on the loss of an industry, or the whole financial system, instead of the issuers' own loss.[15] In this way, contingent capital is only used for managing systemic risk. The issuer is able to raise capital using traditional methods such as rights issues if the issuer gets into trouble due to its idiosyncratic risk. Others argue that an industry-wide trigger may increase the systemic risk instead of decreasing it. If the trigger is at the discretion of the regulators, they might be reluctant to trigger the conversion. If contingent capital is triggered or is near the trigger point, it will convey a very clear adverse message to the market which may lead to overreaction of investors and therefore more

---

[13] Hillion and Vermaelen, "Death Spiral Convertibles," 3-6; MacDonald, "Contingent Capital," 11.
[14] Maes and Schoutens, "Contingent Capital," 7.
[15] Squam Lake Working Party, "An Expedited Resolution," 4.

downward pressure.[16]

(4) While contingent capital may boost the issuer's capital adequacy ratio, it may not be able to solve the liquidity issue, which is normally the direct reason for bankruptcy in a financial crisis. Converting contingent capital instruments such as CoCo bonds to equity will reduce future liquidity requirements such as interest payment. However, unlike other capital raising activities such as rights issue or government bailout, it will not inject liquid assets.

(5) Regulators showed great interest in exploring contingent capital instruments to help prevent the next financial crisis and reduce the usage of taxpayers' money to support system important firms. However, the rule is still vague and under development. If regulators in different regions have different rules for contingent capital, especially on qualification requirements, it could be a difficult situation for global banks and insurers.

(6) Contingent capital might reduce the disciplining power of the debt holders. As Koziol and Lawrenz (2011) showed in a model, when the managers have the discretion of risk taking activities, the bank's probability of financial crisis will be increased by having contingent capital.[17] In addition, there may be an increase in agency cost of equity and therefore the cost of debt due to the reduction in managerial ownership.

3. KEY FEATURES

There are many proposals for using contingent capital to solve the recapitalization issue faced by the shareholders and the too-big-to-fail issue faced by the regulators in a financial crisis. Although contingent capital is seen as a promising candidate to increase the capitalization level and reduce the possibility of government bailout, it does not have a mature market yet. Different opinions of the appropriate design and its complicated features justify more analysis and tests on the market before a full endorsement. This section will discuss the key features of contingent capital that are critical for fulfilling its goal in a practical way. The prominent designs and opinions of contingent capital will also be described.

3.1 Trigger Event: Rule-Based or Discretionary

The designs of contingent capital instruments include the trigger event based on clearly specified rules or at the regulators' or issuers' discretion. For industry-level trigger events, there are different

---

[16] Goldman Sachs, "Contingent Capital," 7, has a discussion about the problems with using a trigger based on regulatory discretion.

[17] Koziol and Lawrenz, "Contingent Convertibles," 18-34.

opinions about the appropriate choices. For institution-level trigger events, there seems to be a mutual preference of using a rule-based event. Contingent capital instruments with a rule-based trigger event are perceived to be more transparent, predictable, and attractive to potential investors. It is also easier to price.

Acharya et al. (2009) think that an industry-level trigger must be rule-based, rather than at the discretion of regulators. With a discretionary feature, the occurrence of trigger would convey severe adverse news to the market, causing a possible downward spiral. In contrast, a rule-based trigger would be well-anticipated and would not have such consequences. In addition, the political pressure on the regulators for the announcement of conversion is not trivial due to its signaling effect. Squam Lake Working Group (2009) has a different opinion for the rule-based industry-level trigger: they are concerned that the aggregate data regulators might use are likely to be imprecise, subject to revisions, and measured with time lags.

Rating agencies require an objective and rule-based trigger event as one of the preconditions for assigning a credit rating. Therefore, a rule-based trigger event is more promising from the perspective of the marketability of contingent capital instruments.

The capital access bond (CAB) proposed by Bolton and Samama (2011) is an exception for institution-level trigger event. The issuer has the option to convert the CAB into equity at the pre-specified price and also has the full discretion on the conversion. Technically speaking, the conversion is still based on rules. The conversion will happen if the option is in the money at maturity. However, as pointed out by the designers, the signaling effects might prevent a decision based solely on the payoff of the option. Not converting the CAB when the conversion price is less than current market price of the stock could be conceived by the market as a higher equity value, which has a positive impact. Or the bank's rational management will be questioned, which has a negative impact.

3.2 Trigger Event: Institution Level and/or Industry Level

Existing proposals of contingent capital have a conversion trigger based on the issuer's financial condition and/or on an industry-wide indicator. The issuer's financial conditions can be indicated by its stock price, capital adequacy ratio, or book value of equity. Industry-wide indicators include aggregated market loss indices and financial industry loss indices. There is no mutual agreement on the type of trigger events to be used among the academic and the industry. Up till now, most existing contingent capital deals are based on an institution-level trigger.

An institution-level trigger event has a focus on the financial condition of the issuer. It is not directly linked to systemic crisis. Therefore, contingent capital with institution-level trigger absorbs loss caused by systemic risk and non-systemic risk. Someone might have concerns about the unnecessary protection for firms taking more than necessary non-systemic risk. Flannery (2009) pointed out that systemically important firms cannot be permitted to fail, as it might cause market turbulence, no matter what the cause is. The prevention from bankruptcy due to non-systemic risk might help protect the incapable managers. Flannery (2009) argued that it is a general corporate governance issue and is not something new brought by contingent capital. Contingent capital can at least protect the taxpayers as intended.

An industry-level trigger event may be more suitable for mitigating systemic risk. With the industry-level trigger in place, contingent capital instruments will be converted only in a systemic crisis. However, it may be hard to implement in an objective way and may convey an adverse message to the market upon conversion. If the industry-level trigger is at the regulators' discretion, due to the signaling effect of conversion, there is political pressure which might delay the triggering and therefore cause more loss. A trigger event that is solely based on the industry-level condition could act as an disincentive for sound risk management, as all the firms are treated the same, no matter how much systemic risk they contribute to the industry.

Some proposals include a dual trigger based on both an institution level condition and an industry level condition. Only when both conditions are true will the conversion be automatically triggered. Examples include Squam Lake Working Group (2009) and McDonald (2011).

3.3 Trigger Event: Based on Book Value or Market Value

Another key element of trigger events is the basis of value measurement. It could be a book value measure which is based on accounting rules or regulatory rules. Or it could be a market value measure determined by investors. Both types of measurement have their own shortcomings but a measure based on market value is preferred by the academic community.

Using a trigger event based on market value might cause the following issues.

(1) Book value of equity is subject to the adjustment of management. Given the current level of complexity of financial institutes, a management team has leeway to move accounting entries off the balance sheet, and therefore the equity under GAAP or IFRS could be manipulated.

(2) Book value of equity is a backward-looking measure for some accounting frameworks such as U.S. GAAP, where historical cost plays an important role in valuation.

(3) The reporting of financial conditions is not continuous. The timing of the trigger that is

based on the financial reports might lag behind the time when recapitalization is needed based on market condition. The timing issue might prevent the contingent capital from being converted to equity before the company goes bankrupt or receives government aid.

Atrigger event based on market value will not have these issues. However, it has its own problems.

(1) The stock price is subject to market manipulation. The impact could be quite significant for contingent capital whose conversion will cause material dilution of shareholders' value.

(2) As pointed out by Flannery (2009), stock price is subject to random pricing errors and therefore random elements in the conversion based on the stock price. The impact is more prominent for the design where conversion price is the same as the then-current market price at the date of the conversion. However, the impact could be dampened if the conversion price is set to be the average of the daily market prices in a certain time period with a fixed length.

Acharya et al. (2009) and Flannery (2009) clearly state in their reports that a market-based trigger is more appropriate due to its timeliness and less exposure to managerial manipulation. However, some existing arrangements have a capital adequacy ratio based trigger event which is a book value measure based on regulatory rules. Lloyds Banking Group's ECNs and Rabobank's senior contingent notes issued in 2009 are real examples. The reasons for choosing a capital adequacy ratio might be the following.

(1) Using a capital adequacy ratio is straightforward regarding the goal of reducing the possibility of default or government bailout, although the failure to fulfill obligation might be caused directly by liquidity issue.

(2) Some arrangements have the level of the trigger well above the minimum requirement, sotimeliness would be less of an issue. The conversion is more likely to happen before the issuer becomes insolvent.

3.4 Conversion Price

Conversion price determines how many shares investors will get if the conversion is triggered. It could be set as:

(1) A fixed value, such as the stock price at the issue date of the contingent capital instrument. Some contracts specify the number of shares to be received upon conversion instead of the conversion price.

(2) The stock price at conversion, which is also known as par conversion.

(3) The stock price at conversion with a floor.

(4) The stock price at conversion with a discount or a premium.

As a variant, the stock price at conversion could be determined as the average of daily stock prices at and before the conversion.

McDonald (2011) provided a comprehensive analysis of different types of conversion price, including fixed share conversion, fixed dollar conversion, par conversion, premium conversion, and discount conversion.[18] Most discussions on how to choose a conversion price are about the implication on the dilution of shareholders' value (Section 4.2), potential price manipulation (Section 3.5), and multiple equilibria of the market equity trigger (Section 3.6).

3.5 Market Manipulation

Market manipulation is one of the major concerns for the effectiveness of contingent capital in reducing systemic risk. The investors of the contingent convertible bonds might short sell the issuers' stock to limit their loss. Short selling is more likely when the stock price drops to a level close to triggering. If it is a conversion at par, there is more incentive to bring down the stock price, as it means more value transferred from shareholders to investors of contingent capital. Anticipating this, existing stakeholders will also sell their holdings to reduce their loss as soon as possible when the stock price is close to the conversion point. This will put extra pressure on the stock price. This phenomenon is known as the death spiral. The death spiral impact of convertible securities is not something entirely new. Hillion and Vermaelen (2001) investigated the death spiral convertibles[19] and found that material loss occurred for the investors.

However, whether the death spiral will continue after the conversion is questionable. If the stock price is too low, there might be less liquidity and a higher bid-ask spread. If the loss is material, investors might want to buy shares instead of selling them with the hope of a recovery or government bailout.

Some adjustment of the conversion price might offset the impact of market manipulation. McDonald (2011) argued that a fixed share premium convertible structure is least exposed to

---

[18] McDonald, "Contingent Capital," 5.

[19] "Death spiral convertibles are privately held convertible securities (preferred stock or debentures) with a conversion price that is set at a discount from the average (or sometimes the minimum) of past stock prices in a look-back period." Hillion and Vermaelen, "Death Spiral Convertibles," 1.

manipulation. The lower the stock price is, the less the value after conversion. A conversion with premium means a further reduction of value after conversion. In fact, investors might influence the stock price in the opposite direction to avoid conversion. However, it is more difficult to push up the stock price if the issuer is in distress.

Another commonly suggested way of preventing market manipulation is to determine the conversion price as the average of the past n days' stock price prior to the conversion. McDonald (2011) and Flannery (2009) suggested that this feature makes it more difficult for manipulation due to the lengthened period of holding short positions. The disadvantage of this structure is that the conversion might be delayed. Squam Lake Working Group (2009) also pointed out another possible manipulation: "*If the stock price falls precipitously during a systemic crisis, management might intentionally violate the trigger and force conversion at a stale price that now looks good to the stockholders.*"[20]

Flannery (2009) and McDonald (2011) also examined the possibility of retiring the contingent capital upon conversion gradually and randomly to avoid a huge gain from price manipulation. As pointed out by Flannery (2009), forbidding holders of contingent capital to short sell the issuer's stock is also a possible solution.

Another potential method of price manipulation is share repurchase. If the conversion is believed to be highly probable to occur, with a material value dilution, the issuer has an incentive to prevent the conversion by putting upward pressure by share repurchase. McDonald (2011) thinks that the impact would be small. If the market thinks that the goal of share repurchase is to avoid conversion, it will have a negative impact on the share price, and so it might not actually happen.

3.6 Multiple Equilibria

One interesting conclusion by Sundaresan and Wang (2011) is that multiple equilibria or no equilibrium may exist for contingent capital with a market trigger.[21] When the stock price is close to the trigger point, there could be different speculation on the occurrence of conversion. The equilibrium stock price near conversion could be different depending on whether or not the conversion will happen. Two equilibrium prices are possible: (1) the equilibrium price which is above the conversion price, assuming there will be no conversion; and (2) the equilibrium price which is below the conversion price, assuming there will not be conversion.[22] Sundaresan and Wang (2011) point out that multiple equilibria are caused by the value transfer between shareholders and

[20] Squam Lake Working Group, "An Expedited Resolution," 5.
[21] Market trigger is a trigger on market value of equity.
[22] Mathematical deduction and numerical examples can be found in Sundaresan and Wang, "On the Design," 9-12.

contingent capital holders. The value transfer upon conversion will cause a sudden change in the stock price, the source of the difference in multiple equilibrium prices. However, this would not be an issue if the conversion is made at par,[23]

If the conversion is not timely, the stock price can fall below the trigger point before the conversion of a CoCo bond. Sundaresan and Wang (2011) argue that there is no equilibrium price in this situation. If the conversion is believed to happen, the stock price will fall below the conversion price and the conversion will happen. If the conversion is believed not to happen, the stock price will be higher than the conversion price. Both scenarios are not consistent with the assumption that the stock price could fall below conversion price without a conversion.

Prescott (2011) illustrated that a trigger on the market of equity could potentially cause multiple equilibria and nonexistence of the equilibrium. Based on the market experiment data and empirical evidence, it is argued that a market equity trigger "*made prices and allocations less efficient and led to numerous conversion errors.*"[24]

Theoretically, the lack of a unique price indicates an unstable market and a high exposure to price manipulation. The market price would therefore not represent its true economic value and might cause market inefficiency. However, the nonexistence of equilibrium happens when the stock price is near the trigger point. Given that contingent capital is used to deal with financial distress, the stock price is volatile during that period even without the presence of contingent capital. At that time, the equity market is in an unstable situation and the stock price is highly driven by the psychological factors of investors and sensitive to any new information. Multiple equilibria or the absence of equilibrium caused by contingent capital may make it more complicated, but not necessary worse. The market expectation of the occurrence of conversion will be determined by market participants, just as participants determine the stock price. When the expectation changes, the stock price can suddenly jump. A jump in stock price is not abnormal when the company is in financial distress. Other investor behaviors may also dominate the movement of stock price and push it to a certain equilibrium price. For example, the short selling behavior as discussed in Section 3.5 may move the stock price quickly until at or below the conversion price.[25] It would eliminate one of the two possible equilibrium prices. Therefore, multiple equilibria may not cause chaos in practice.

---

[23] A conversion at par indicates that there is no value transfer between shareholders and contingent capital investors.

[24] Prescott, "Contingent Capital: The Trigger Problem," 15.

[25] There is incentive for short selling behavior as long as the conversion price is linked to the then-current equity price even if the conversion is not at par.

Contingent capital instruments with trigger events based on book value, regulatory solvency ratio, and industry level loss are less exposed to the issue of multiple equilibrium. Although some trigger events are highly correlated with stock price and there could be value transfer between contingent capital holders and shareholders, the complexity and the uncertainty of the accounting and regulatory rules make it difficult to speculate based on the occurrence of conversion and the equilibrium price. Importantly, note that placing the trigger on firm value is not equivalent to placing the trigger on the market equity ratio. A trigger based on firm value is not affected by the multiple equilibrium issue.

3.7 Proposals

There have been many proposals about the appropriate design of contingent capital since 2008 from the academic community. Some of them are described below to illustrate the variety, complexity, and the ongoing development of contingent capital market.

Squam Lake Working Group (2009) suggested a conversion from debt to equity if two conditions are met. The first condition is an industry-level event such as a declaration by regulators that the financial system is suffering from a systemic crisis. The second is an institution-based event such as a violation of covenants in the hybrid-security contract. A promising candidate of the covenant is the capital adequacy ratio (Bank's Tier 1 Capital/risk adjusted assets).

McDonald (2011) analyzed a dual trigger design based on the firm's stock price and the value of a financial institutions index. This structure potentially protects financial firms during a crisis, when all are performing badly, but during normal times it allows a bank with bad performance to go bankrupt.

Kashyap et al. (2008) proposed using capital insurance that would "*transfer more capital onto the balance sheets of banking firms in those states when aggregate bank capital is, from a social point of view, particularly scarce*."[26] It is an insurance contract, not like a contingent convertible bond. The trigger of insurance payoff is based on the capital loss of the total banking industry. The insurance payment can help strengthen the solvency position and provide liquidity.

Flannery (2009) proposed "*Contingent capital certificates" (CCC) that "would be issued as debt obligations, but would convert into common stock if the issuer's capital ratio fell below some critical, pre-specified value.*" It is suggested to be applied to systemically important firms. However, the condition of conversion and its specification are complicated compared to other proposals. Flanery's key features are quoted

[26] Kashyap et al., "Rethinking Capital Regulation," 452.

below.

a. *A large financial firm must maintain enough common equity that its default is very unlikely.*

b. *This common equity can satisfy either of two requirements:*

- *Common equity with a market value exceeding 6% of some asset or risk aggregate.*
- *Common equity with a market value exceeding 4% of total assets, provided it also has outstanding subordinated (CCC) debt that converts into shares if the firm's equity market value falls below 4% of total assets. The subordinated debt must be at least 4% of total assets.*

c. *The CCC will convert on the day after the issuer's common shares' market value falls below 4% of total assets.*

d. *Enough CCC will convert to return the issuers' common equity market value to 5% of its on-book total assets.*

e. *The face value of converted debt will purchase a number of common shares implied by the market price of common equity on the day of the conversion.*

f. *Converted CCC must be replaced in the capital structure promptly.*

g. *The CCC debt converts automatically – no option. If the firm is insolvent when conversion is triggered (e.g., because of a jump in asset values), the debt covenants must specify a conversion price that wipes out the previous shareholders.*

h. *CCC cannot be owned by systemically important firms for their own account.*

i. *The CCC that will be converted need to follow some selection rules such as retiring the shortest maturity, random selection, and according to the seniorities of the CCC.*[27]

Bolton and Samama (2011) proposed the capital access bond (CAB), "*which gives the issuer of the bond the unconstrained right to exercise the option to repay the bond in stock at any given time during the life of the bond. It is effectively an option to issue equity at a prespecified price, with the added feature that the writer of the option puts up collateral to guarantee that it is able to fund the purchase of new equity should the buyer of the put option choose to exercise the option.*"[28] This is quite similar to the reverse convertible bond that may be converted into stock or its cash equivalent at maturity or at some triggering event. CAB has two embedded options: a call option on the bond and a put option on the shares. Unlike the traditional convertible bond, both options are owned by the issuer. The idea of issuing multiple CABs with

[27] Flannery, "Stabilizing Large," 9-11.
[28] Bolton and Samama, "Capital Access Bonds," 10.

different strike price is also discussed to deal with all kinds of contingencies in a crisis.

Calomiris and Herring (2011) proposed a contingent capital instrument with a quasi market value based equity-ratio trigger designed to smooth the impact of the fluctuations in share prices. It is calculated as the 90-day moving average of the ratio of the market value of equity to the sum of the market value of equity and the face value of debt. The bond is not expected to be converted but will promote more efficient corporate governance. The trigger event is designed to be less exposed to manipulation. The conversion price is also set to have a material dilution of shareholders' value.

The diversity of the proposals listed indicates that contingent capital is still under development both in theory and in practice. The surveys of the existing literature on contingent capital are not rare. Calomiris and Herring (2011)[29] did a comparison of different designs regarding the amount of CoCo bonds required to be issued, the trigger event, and the term of conversion. Cooley et al. (2010)[30] summarized the types, the trigger events, whether they are equity based or credit based, whether they have market value trigger or book value trigger, and the drawbacks.

3.8 Post Conversion

Ownership

There are concerns that the conversion of contingent capital gets the distressed firms out of trouble and at the same time keeps the incompetent management team. As Flannery (2009) pointed out, it is a general corporate governance issue and is not something new brought by contingent capital. Collender et al. (2010) mentioned that contingent capital could "*require the replacement of or votes to replace management and the board of directors*"[31] if a certain amount of contingent capital has been converted. There are concerns that contingent capital might dampen the disciplining power of debt holders. Allowing the replacement of management and board of directors might be a good idea, as this threat may discourage the management to take aggressive actions trying to recover their loss without caring for the downside risk. Such kind of terms written in the contingent capital contract would make it a complete contract[32] where the probability of default is shown by Koziol and Lawrenz (2011) is reduced.

New Issuance

After contingent capital is converted, it is hoped that the firm will return to a healthy condition.

---

[29] Calomiris and Herring, "Why and How," 42-46.
[30] Cooley et al., "Regulating Wall Street," 168-174.
[31] Collener et al, "Automatic Recapitalization," 13.
[32] Koziol and Lawrenz, "Contingent Convertibles," 14-18.

However, the firm will have no or less automatic bail-in securities after conversion. Therefore, it is necessary to issue new contingent capital instruments in a short period to regain the protection.

3.9 Contingent Capital versus Other Risk Transfer Instruments

The insurance industry has a history of transferring risks through traditional reinsurance arrangements and alternative risk transfer instruments. The majority of the risks involved are insurance risks such as catastrophe risk, mortality risk, longevity risk, and lapse risk. Others usually help transfer all the risks of the written business, including financial risks to a third party. This section compares those existing risk transfer instruments to the relatively new contingent capital instruments. They are all utilized for risk mitigation but their focuses are quite different.

Traditional Reinsurance Arrangement

Reinsurance has been used by insurance companies to transfer undesired risks, stabilize their claim experience, increase their capacity of writing new business, and improve the efficiency of capital usage. Depending on the type of reinsurance arrangement, specific risks or all risks of insurance business are transferred from the primary insurer to the reinsurer. It is quite different from contingent capital in the following aspects.

(1) The loss-absorbing capacity of the reinsures is less than the potential market for contingent capital, the whole capital market.

(2) Reinsurance deals with the risk on the liability side, while contingent capital works as a buffer for the risks from both the asset side and the liability side.

(3) The primary insurer is exposed to counterparty risk, since the reinsurer may fail to pay the reinsurance claim. That is not the case for the issuer of contingent capital, since the price is paid at the issue date.

(4) Contingent capital provides funds under distressed situations when the conversion option is exercised, while reinsurance is used as a general risk transfer channel. From the perspective of maximizing the risk adjusted return on capital, contingent capital is probably a better choice given the relatively high cost of buying reinsurance protection.

Insurance Derivatives

There are several types of non-traditional financial instruments that have payments contingent on a certain insurance event, loss, or experience. Those insurance derivatives are used to transfer insurance risk to reinsurers or general investors.

(1) Industry loss warranties (ILWs) deal with the loss caused by insurance events such as hurricanes, windstorms, and earthquakes. The owner of the contract will get paid a specified amount if the industry loss caused by the disaster exceeds the trigger level.

(2) Catastrophe bonds are sold to reduce the exposure to catastrophe risk. The investor will get a rich coupon payment but will lose the coupons and/or the principle once a catastrophe event occurs.

(3) Longevity bonds have the amount of coupon payment linked to the number of survivors for a chosen population cohort. It is often used to hedge the risk that one outlives his/her savings. Insurance companies buy longevity bonds to protect them against the risk that the mortality experience is better than what is assumed when pricing life annuities.

In contrast, contingent capital focuses on the financial risks, especially the systemic risk. In addition, the conversion is expected to happen only under financial stress. The writing down of liability or the conversion to equity helps strengthen the capital position. The insurance derivatives are used to limit the loss no matter whether the buyer is in financial trouble or not.

Sidecar

A sidecar, an arrangement that allows the investors to get the return and take the risk of insurance business, can be used to enhance the insurers' ability to take risk. Normally, a special-purpose vehicle (SPV) needs to be established. The insurer pays a premium to the SPV and the investors deposit money to the SPV to cover the claims by policyholders. It is similar to quota-share reinsurance that transfers part of the business and risk to the reinsurer. The insurer transfers the written business to the investors through the sidecar. However, the entire risk is not transferred. The loss of the investors is limited to the funds that are put in as required. When the realized experience is worse than what the fund can cover, the insurer has to bear the remaining loss. However, the required fund value is normally high enough so that the chance of excess loss is slim.

A sidecar helps reduce the risk exposure and the reserve that is required to support the transferred insurance business. Both sidecars and contingent capital have a positive impact on the capital position, but they are meant to meet different challenges.

(1) The return and risk is transferred to outside investors for certain written business through the sidecar. However, contingent capital is normally linked with the risk of the whole portfolio, including the retained business as well.

(2) The sidecar is normally arranged to optimize the usage and efficiency of the capital and increase the risk-taking capability. When an insurer or reinsurer finds a profitable investment

opportunity but does not have enough capital to invest, it may consider using a sidecar to release some used capital to take the opportunity. In contrast, contingent capital is deployed to increase the loss absorbing capability and reduce the risk of insolvency. Therefore, a sidecar is used to prepare for taking more risks while contingent capital is used to prepare for future distressed situations.

Catastrophe Equity Put

Catastrophe equity put (CEP) is a special type of contingent capital with a focus on the catastrophe risk. Given the material impact of catastrophe events on the loss and stock price of insurance companies, CEP appears to be an effective tool to mitigate the risk by flooring the equity value. Compared to standard equity put option contracts, it requires two conditions to be met contemporarily before exercising the option.

(1) The equity price drops below the exercise price

(2) The catastrophe loss exceeds the prespecified amount

Whereas other contingent capital instruments normally have triggers related to the general financial health condition of the issuer or the financial industry, CEP is used for mitigating catastrophe risk only, not financial systemic risk or any other kinds of risk. CEP and other types of contingent capital are not exclusive but complementary.

Line of Credit

A line of credit may be more appropriately termed a risk mitigation tool rather than a risk transfer instrument. It is a source of financing provided to credit-worthy companies or individuals by banks. The borrower can use the fund and need to pay interests and other related fees. Insurance companies, especially when publicly listed, often use line of credit to increase its available liquidity source and therefore help reduce the exposure to liquidity risk.

Lines of credit and contingent capital are designed for different purposes. Lines of credit only allow the company to borrow extra cash at a cost. They do not change the composition and amount of the surplus account. When there is a capital inadequacy problem, borrowing money will not be helpful, as cash borrowed will be reflected on the liability side as well. Available capital will hardly change. On the other hand, the conversion of contingent capital will lead to a direct liability written down or an increase in the available capital. These two sources of financing deal with different kinds of risk and therefore they behave differently in many aspects.

4. EFFECTIVENESS AND POTENTIAL IMPACT

4.1 The Impact on Systemic Risk

Systemic risk is the risk that the entire financial system may collapse. Due to the interdependence within the financial system, the failure of one or more systemically important firms can lead to the crash of the entire system. Contingent capital is believed to reduce the systemic risk and default probability when compared to pure debt instruments. Issuing contingent capital without reducing the common equity provides additional capital which would certainly reduce the chance of bankruptcy. However, substituting common equity with contingent capital to meet capital requirement might have an uncertain impact on systemic risk due to the uncertainty of the conversion.

The effect of reducing systemic risk was shown by Hilscher and Raviv (2011) analytically via a quantitative model. Based on the model setup, contingent capital financing can reduce the default probability compared to subordinate bond financing. In addition, the risk-taking incentive becomes less if the conversion price indicates a certain level of value dilution for the existing shareholders.

Although a properly designed contingent capital may reduce systemic risk, it is not expected to be the entire solution for the too-big-to-fail issue. McDonald (2011) emphasized that the proposed dual trigger contingent capital instrument reduces a firm's debt load but is not used to address the too-big-to-fail issue. Regulators need to "*proactively monitor the management and performance of financial institutions. Contingent capital is thus a backstop for regulatory failures or unforeseen market events, not a regulatory substitute.*"[33]

Acharya et al. (2009) also pointed out that contingent capital is not enough for eliminating systemic risk. Even when the issuers in trouble remain solvent due to the timely conversion from debt to equity, the firms are still exposed to liquidity risk which is normally the direct cause of going bankruptcy. Counterparty risk also exists, although it is expected to be lower if every firm's contribution to systemic risk is reduced by issuing contingent capital. The loss in excess of the value of the equity and contingent capital for too-big-to-fail firms is still likely to be protected by the government. Therefore, there is still an incentive to take risk above a firm's capability. Acharya et al. (2009) concluded that "*an explicit fee … charged to banks in good times based on their expected losses and their systemic risk contributions*" is necessary for rectifying the moral hazard due to the implicit guarantee by the government.

---

[33] McDonald, "Contingent Capital," 2.

4.2 Dilution of Shareholder Value

When the conversion price is set to be equal to the market price, there is material dilution of shareholder's value on the conversion of the CoCo bond. A fixed or floored conversion price will limit the dilution. With the threat of value dilution upon conversion, existing shareholders are discouraged from taking excessive risk above their risk tolerance. However, by fixing the conversion price to be the stock price at the issue date of the CoCo bonds, the magnitude of dilution of the shareholder's value is floored. It might not provide enough incentive for disciplined strategic planning. Calomiris and Herring (2011) mentioned that in the interests of creditors and regulators who are more risk averse, a dilution of shareholders' value is critical for a more stringent control on risk-taking activities.

Investors' behavior could also have an impact on the value transfer from equity holders to the contingent capital investors. If the contingent capital instrument has a conversion price equal to the then-current stock price, contingent bond holders will probably short sell the stock near the trigger point as a sharp drop in stock price is favorable. Low stock price means a high value transfer. Shareholders who foresee this behavior will also sell their shares as soon as possible, hoping to get a higher price than the conversion price.

4.3 Capital Admittance and Accounting Treatment

The regulation and financial reporting rules have been evolving and it is not crystal clear at this point what the final decision will be for the treatment of contingent capital instruments.

<u>Capital Admittance</u>

Under the current capital requirements directive (CRD II) of European Union, instruments that absorb losses on a going-concern basis and that must be converted to core Tier 1 capital are regarded as equity capital, capped by 50% of the core Tier 1 capital. Contingent capital instruments with those features are likely to be classified as Tier 2 capital under current framework. According to the consultation paper (CRD IV) issued in 2010, European Commission will consider the potential need for all non-core Tier 1 instruments to have a mandatory principal write-down or conversion feature, the potential triggers for conversion, and alternative mechanisms and triggers of contingent capital.[34] Therefore, it is possible that certain types of contingent capital instruments could qualify for non-core Tier 1 capital.

As written in a future rule about the additional loss absorbency requirements for global

[34]COMMISSION SERVICES STAFF WORKING DOCUMENT, "POSSIBLE FURTHER CHANGES," 18-19&23.

systematically important banks (GSIB) released by Basel Committee on Banking Supervision (BCBS) in late 2011, contingent capital will have to be converted to Common Equity Tier 1 when the Common Equity Tier 1 falls below 7% of risk adjusted assets. It also sets a cap on the new shares and requires a full authorization from the issuers for an immediate conversion.

Under the Solvency II framework, contingent capital with appropriate feature can be used to meet the solvency capital requirement (SCR), but not the minimum capital requirement (MCR). However, the amounts of ancillary own fund (AOF) items and the amount for each AOF item are subject to supervisory approval.[35]

In its final advisory on Non-Viability Contingent Capital, the Office of the Superintendent of Financial Institutions Canada (OSFI) requires that the regulatory capital of all federally regulated deposit-taking institutions (DTIs) must have a loss-absorbing quality when the DTI fails. All the non-common Tier 1 and Tier 2 capital must satisfy the requirements for non-viability contingent capital (NVCC). However, the trigger event of NVCC is at the discretion of the regulators. Contingent capital instruments that meet the requirements are likely to qualify for either non-common Tier 1 or non-common Tier 2 capital.

In the United States, the treatment of contingent capital is not clear, both for the banking industry and the insurance industry.

Accounting Treatment

There is, as of yet, no guidance issued for contingent capital regarding the accounting treatment. According to IAS 32, under IFRS, convertible bonds need to be presented as two components on the issuer's balance sheet:

(1) A financial liability whose value is determined by measuring the fair value of a similar liability that does not have the conversion option, and

(2) An equity instrument whose value is determined as the fair value of the option to convert the instrument into ordinary shares.

Unlike convertible bonds, the conversion option of contingent capital instruments is rule-based or at the discretion of the issuer or the regulators, not the investors. It is likely that contingent capital instrument will be presented as two components under IFRS in the same way, except that the value of the equity instrument depends on a different option.

[35] "Directive 2009/138/EC," 48.

In IAS 32, it is mentioned that "*Classification of the liability and equity components of a convertible instrument is not revised as a result of a change in the likelihood that a conversion option will be exercised, even when exercise of the option may appear to have become economically advantageous to some holders.*"[36] The arguments are that the holders may not always act in the way that might be expected and the likelihood of conversion will change from time to time. Certainly, the same arguments also apply to contingent capital. It is likely that under IFRS, once the classification of the liability and equity components is determined at issuance, it will be kept the same until it is converted.

U.S. GAAP has different treatment for convertible bonds from IFRS. According to FAS 133, for the issuer convertible bonds without the cash settlement option should not be separated into two components, as its stock price is closely related to the convertible bonds. On the other hand, for the investors, the embedded conversion option needs to be separated from the debt component under U.S. GAAP. Contingent capital might get different treatment given that regulators consider it as an automatic recapitalization tool to mitigate systemic risk. The equity credits might be allowed to be reflected before conversion. However, there is no clear rule from FASB about contingent capital.

The impact on earnings volatility depends on the financial condition of the issuer or the financial industry. If the contingent capital is treated as debt before conversion and equity after conversion, the earnings volatility of the issuer before conversion will be the same as that for issuing traditional debt. However, upon conversion, a write-down of the issuer's liability will certainly reduce the loss and therefore the earnings volatility. If the conversion option is separated from the host contract, the earnings volatility of the issuer before conversion tends to be more volatile compared to issuing a traditional bond. The mark-to-market value of the conversion option will be a major contributor to the volatility.

4.4 Tax Deductibility

One of the key benefits of issuing contingent capital is the tax treatment which it is expected to receive. The interest payments of contingent capital are tax deductible considering the fact that contingent capital behaves like a debt instrument before conversion. Contingent capital and equity are still quite different given that contingent capital has limited upside while equity has unlimited upside. In addition, the contingent capital is brought in to reduce systemic risk and the cost of government bailout in a financial crisis. Because contingent capital has a combination of limited upside and potential loss absorbency, it makes more sense for it to be tax deductible.

Bolton and Samama (2011) mentioned that Lloyds Banking Group's ECNs and Rabobank's

[36] European Commission, "IAS 32," 10.

Senior Contingent Notes are tax deductible.[37] Although both of the contingent capital instruments were issued since the financial crisis in 2008, the two issuers are under different financial conditions at issuance. Lloyds Banking Group had been bailed out by the government while Rabobank was in good financial conditions and the Senior Contingent Notes were oversubscribed. This indicates the likely treatment that contingent capital might receive, both in good times and in bad times.

Bolton and Samama (2011) also mention a possibility that the option premium as part of the coupon payments is not tax deductible and the movement of the fair value of the embedded option needs to be taxed as an income. Its complexity might make this unlikely. However, the IFRS is likely to require the debt component and the equity component to be separately reported under fair value basis. Therefore, there may still be a chance that the conversion option will be treated differently from the pure debt component.

4.5 Disclosure Requirement

Due to the complexity of contingent capital, transparency is the key to its marketability. Rating agencies have expressed their positive view on this desired feature. Investors would also prefer securities with sufficient information about the issuer. Vagueness means higher required return and lower price of contingent capital instruments which will increase the cost of financing. The report publicized by Goldman Sachs emphasized more standardized bank disclosure as an important factor to make the objective triggers more credible.[38]

Not only is the disclosure of the contingent capital instrument itself necessary for transparency, butmore detailed and timely disclosure of financial condition, business plan, and risk appetite is also needed for the success of contingent capital. The issuer needs to provide enough information to the potential investors so that they can assess the risk and return of the instruments. Without doing this, investors will not be able to make an informed decision and may be reluctant to invest.

4.6 Counterparty Risk Assessment

Contingent capital instruments change the issuer's capital structure and risk-absorbing capability. For the issuer, there is no counterparty risk, as the price is paid at issue. However, the default risk profile of the issuer is changed when contingent capital is issued. Therefore, the counterparty risk of holding contingent capital and the ordinary fixed income securities without conversion options will be affected. In addition, the expected outcome of conversion may not be realized due to the behaviors of the investors and the issuer, as discussed in Section 3.5. This makes the counterparty

---

[37] Bolton and Samama, "Capital Access Bonds," 31-32.
[38] Goldman Sachs, "Contingent Capital," 4-8.

risk analysis even more complicated.

One straightforward approach is to rely on the credit rating given by the rating agencies. However, not all the contingent capital will be rated by rating agencies. The contingent capital instrument has to first meet the standard set out by the rating agencies. In addition, the credit rating may be unavailable or outdated when you need it to make the decision, especially in a stressed situation when the conversion is likely to happen soon.

An alternative approach that also relies on the market available information is the credit default swap spread. The impact of contingent capital on the default risk is expected to be reflected in the CDS spread. However, this information may not be available for certain issuers and is only available after the contingent capital instruments are issued.

Given the drawbacks of the approaches above, sometimes it is necessary to analyze the relationship between contingent capital and default risk in order to quantify the counterparty risk. Some papers have examples of quantifying the impact on default risk. Hilscher and Raviv (2011) showed analytically that contingent convertible bond financing can reduce the default probability compared to subordinated bond financing. The underlying logic is straightforward. The issuer will default if it cannot pay the coupon or redemption value of the subordinated bonds. However, that is not the case if a CoCo bond is used instead. The CoCo bond will be converted before the firm fails to pay the coupon and redemption value. Clearly, this argument is based on the assumption that the risk-taking behavior is the same in those two circumstances. If the issuer thinks that the introduction of CoCo bonds allows for a higher risk tolerance, the default risk could be higher due to the lost of the disciplinary power of debt instruments.

5. PRICING, VALUATION, AND RISK ASSESSMENT

5.1 Pricing Models

Contingent capital, as an innovative hybrid security, contains elements of both the debt and the equity. Generally, there are two types of models to price contingent capital.

(1) The first type is based on the Merton (1974) model and the Black Scholes (1973) model. Sometimes it is called structural model. In the Merton model, the shareholder's value is considered as a call option on the firm's value with an exercise price equal to the value of the debt. By translating the trigger event into an equivalent value of the firm, Merton's model can be revised to model the probability of conversion using the equity call option model with a revised exercise price. Existing literatures about pricing contingent capital have more

focus on this approach.

(2) The second type is based on Duffie and Singleton (1999), which models defaults and values corporate bonds through the term structure of interest rate. Sometimes it is called reduced-form model. Unlike the Merton model, it does not explicitly consider the debt structure and the value of the firm at default, or, in other words, the exercise price of the call option. It models the default probability with a hazard rate influenced by exogenous market factors that are closely correlated with the firm value. To adjust for the feature of contingent capital, default intensity needs to be changed to trigger intensity which is the hazard function for the conversion. With Duffie and Singleton's approach, both default intensity and loss ratio can be set as a function of an exogenous variable, such as the stock price. Therefore, stakeholder behavior can be explicitly incorporated in the model. The model is also capable of including jumps for discontinuous information, such as capital rule changes in the diffusion process.

An extension of the first type is to include discontinuous jumps in modeling asset price. The classic Merton model assumes that the firm value follows the pattern of Geometric Brownian Motion. Geometric Brownian Motion is not good at explaining material value change over a short time period, which is a normal phenomenon for stock price. In addition, some contingent capital has its trigger event subject to ad hoc factors. As an example, for contingent capital with a trigger event based on capital adequacy ratio, discontinuous change in the value of the conversion option will happen when there is a change in capital rules, or business strategy, or when updated information is released, such as the capital position in a quarterly financial report. A compound Poisson Process is a common choice for the jump part in jump diffusion models used in the area of finance and risk management. It assumes an exponential distribution for the waiting time between jumps. The jump size follows a specified distribution itself. Sometimes, jump diffusion models are coupled with stochastic volatility, such as the Heston Model. This extra layer of flexibility can be used to take into account some stakeholder behavior such as shorting stocks before conversion.

This section discusses some of the models for contingent capital in existing literatures and the possible improvement.

<u>CoCo Bonds</u>

Spiegeleer and Schoutens (2011) suggests a credit derivatives approach which determines the value of the credit spread on contingent convertible bonds by (1-Recovery Rate) × Trigger Intensity. Recovery rate is the ratio of the share price at conversion to the conversion price. The trigger intensity is associated with the probability of triggering. Realizing the difficulty in modeling the capital adequacy ratio based on regulatory rules or the financial ratio based on accounting rules, the trigger event needs to be translated to an equivalent event such as the stock price dropping below a barrier. Utilizing the well-established digital barrier option pricing method, the probability of a trigger and the trigger intensity can be calculated. As the authors realized and mentioned in the paper, the shortcoming of a credit derivatives approach is that when the share price on the trigger date is very close to the conversion price, the recovery rate is almost one which would indicate that a contingent convertible bond is risk-free. In addition, a credit derivative approach does not consider the possibility that the coupon payment will cease if default event or a conversion happens. A low credit rating of CoCo bonds would indicate a high coupon rate, which has a big impact on the bond price. Equity derivative approach was suggested to take those missing factors into consideration. The CoCo bond can be valued as below.

(1) Plain Vanilla Corporate Bond

\+ (2) Knock-In Forwards between spot price and conversion price[39]

\- (3) Down-and-in cash-or-nothing binary option on the coupon payments

Both the knock-in forwards and down-and-out cash-or-nothing binary option have stock as their underlying asset. Therefore, it is critical to determine the appropriate barrier for the stock price that is equivalent to the occurrence of the trigger event, at least approximately. However, how to translate the trigger event into an equivalent stock price and how to determine the trigger intensity is vague in Spiegeleer and Schoutens (2011). The approaches also neglect the possibility that without the conversion, the financial institute could go bankrupt directly, especially if the checking of the trigger event is not frequent.

A way of deriving the barrier for the stock price is to use the market price of existing deals. The authors illustrate the process of deriving an implicit barrier based on Lloyd's deals. If a similar CoCo bond exists in the market, the implicit barrier approach might be used as a reference point. However, there are several issues with this approach.

(1) Even if the CoCo bonds have the same features, the financial condition, business profile, and the strategic plan could be quite different. Those factors have a big impact on the probability of conversion.

(2) Until the CoCo bond market develops to be a much bigger one, lack of liquidity and transparency would impede us from adopting this approach.

But it is still valuable as a reference check for the price of the new issuance.

---

[39] The knock-in forward will be effective if the trigger event for the CoCo bond happens.

Takahashi et al. (2001) propose an approach of pricing convertible bond based on the Duffie and Singleton (1999) model for default risk. They suggest integrating the default risk and stock price by including the default intensity in the stock price diffusion process. It looks like a very promising candidate for valuing contingent convertible bond, as the trigger intensity could be highly correlated with the stock price, depending on the design of the CoCo bonds.

In a report by FitchSolutions, a structural approach which is more complicated than Merton's model is used to value a CoCo bond. The analytical first passage time model specifies a time dependent threshold which allows more flexibility in the calibration and pricing process. This model was used to value bonds with credit risk. It is modified to value contingent capital, considering both the probability of conversion and the probability of default. The paper also proposed a way to estimate the regulatory capital position based on the leverage ratio.

For contingent capital securities with a dual trigger design, the model could be more complicated. McDonald (2011) provided a pricing example of a dual trigger convertible bond based on the assumption that stock price and systemwide index follows correlated Geometric Brownian Processes with mean reversion. A simulation method is used to calculate the price of contingent capital securities. Historical volatility is used for the volatility parameters and the correlation. In practice, those assumptions might need forward-looking elements. In addition, a linear correlation might be too aggressive an assumption. As the conversion normally happens in a tail event, stock price and index value tend to move at a closer pace than in normal circumstances.

Bolton and Samama (2011) proposed the valuation methods of capital access bond (CAB). CAB is similar to traditional convertible bonds with the exception that the owner of the conversion option is the issuer, not the investor. The issuer also has the option to call the bond before maturity. The option premium for covering the cost of conversion option and call option can be determined using standard option valuation formula. However, the probability of default should be considered. The option premium will be unpaid if the issuer goes bankrupt. The conversion option premium before adjustment also needs to reflect a certain degree of default risk. This is because it is possible that the issuer may go bankrupt before the CAB is converted. Therefore, using an adjusted option price formula is not ideal for valuing CAB. A trinomial tree model was also illustrated by Bolton and Samama (2011). It is more rigorous than using standard option pricing formula as it models the dynamic process of stock price and the conversion.

Each type of the pricing models has its weakness. In structural models, one of the most critical parts is to determine the barrier for the stock price. For trigger events that are directly based on the stock price, it would not be an issue. However, for trigger events that are based on a capital adequacy ratio, it would be very difficult to have a reasonable estimate due to the managerial discretion, the regulatory change, and the time lag between the release of the capital report and the change of stock market.

Another thing to keep in mind when using structural models for pricing contingent capital is that the volatility parameters should be calibrated to appropriate target. Due to the non-flat volatility term structure, the implied volatility of an at-the-money equity put option could be quite different from that of an out-of-money equity put option. Given the level of the trigger point and financial condition of the issuer, an equity put option with a strike price close to expected equity price near conversion should be chosen as the calibration target. However, there may not be enough market information for out-of-money equity options in practice. It could be a major limitation of using structural models in such a situation.

On the other hand, reduced form models need to calibrate the hazard rate of conversion, which poses a big challenge as well. Selecting an appropriate state variable that the hazard rate has a high dependence on and figuring out the relationship between them is not an easy task.

In addition, without explicitly considering the cause and effect, some facts about the issuer, such as its capital structure, may not be appropriately priced in using reduced-form models. It may have a material impact on the price, considering that most of the issuers are large financial institutions with a diverse business and risk profile.

There are some areas that need further research. The price of contingent capital is sensitive to those factors.

(1) Existing models often lack a framework that can explicitly quantify the impact of the stakeholders' behavior, such as the manipulation of the stock price (Section 3.5) and the multiple equilibria issue (Section 3.6).

(2) The conversion event, similar to the default event, deals with the tail risk where market data normally are too sparse to be used for a credible calibration.

(3) Depending on the design of contingent capital, it is possible that default can happen before the conversion option is exercised. It has not been explicitly and fully addressed in existing pricing models.

(4) The impact of the issuer's debt structure on the price of contingent capital needs to be incorporated in the pricing model at a more granular level considering different seniorities of the debt.

(5) New issuance of contingent capital may have an impact on the equity value due to the potential value transfer between shareholder and debt holder and the change in risk taking capability. How to incorporate this kind of change in the stock price in the pricing framework deserves further study.

Catastrophe Equity Put

Catastrophe equity put (CEP) has been used by the insurance industry to transfer the negative impact of catastrophe event on its capital position. It gives the issuer the right to sell stocks to investors at a fixed price once the catastrophe loss exceeds the specified limit. Unlike CoCo bonds, where all the risks could lead to the occurrence of the trigger event, CEP depends on the joint movement of catastrophe losses and stock prices. In other words, the trigger event and the stock price are considered to be highly correlated for CoCo bonds, but that is not the case for CEP. Therefore, it is important to model the relationship between losses and stock price.

Cox and Pedersen (2004) introduced a framework where asset price follows the geometric Brownian model with additional downward jumps when there is a catastrophe event. The jump size is static regardless of the size of the catastrophe loss.

Jaimungal and Wang (2006) generalized the model introduced by Cox et al. (2004) with a stochastic interest rate and the downward jump size depending on the total loss. Unlike Cox et al. (2004), the losses are assumed to follow a compound Poisson process. Jaimungal and Wang (2006) also pointed out that the counterparty risk is not explicitly modeled and the homogenous Poisson process for the catastrophe events is not appropriate for risks with seasonality.

Lin and Chang (2007) made a further step. In the context of catastrophe losses, a constant expected arrival time is not an appropriate assumption. Instead of using a Poisson process, Lin and Chang (2007) proposed the Markov Modulated Poisson Process to model the arrival process of catastrophe events. The arrival process is assumed to follow a homogenous Markov chain which determines the state of the Poisson process. In different states, the arrival rate of catastrophe events could be quite different.

There are two practical issues to consider when using those models.

(1) Both the compound Poisson process and Markov Modulated Poisson Process require the calculation of the cumulative distribution function of the n-fold convolution of losses in the

valuation formula. Except for a few distribution types of the loss size,[40] there is no closed-form formula for calculating aggregate loss. However, there are other methods to compute the n-fold convolution, such as the numerical method and the method of moments.[41]. But it requires either a loss of accuracy or a higher cost of computing resources and time.

(2) The calibration of the loss size model is not an easy task. Historical data of the catastrophe loss and its resulting share price drop are helpful. But a clear picture of the current and future catastrophe risk exposure is necessary to get an up-to-date assumption for the distribution of the downward jump size.

A Stochastic Approach

Some of the pricing models described are analytically tractable and some are not. In light of more powerful computational capabilities, a stochastic approach is another feasible approach. By specifying the models for interest rate, asset price, and catastrophic loss, one can simulate the asset price, the exercise of the conversion option, and the generated cash flows. The value of contingent capital can then be estimated by taking the average of the discounted values for all scenarios. There are two types of stochastic scenarios: risk-neutral scenarios and real-world scenarios. Market-consistent risk-neutral valuation uses risk-neutral scenarios as discount rates while adjusting the probability to match the average discount value with market value. Market-consistent real-world scenarios use the sum of risk-free rate and implied-risk premium as a discount rate. Theoretically, both risk-neutral scenarios and real-world scenarios can be used for market consistent valuation. But risk-neutral scenarios are preferred, as they are more practical for calibration. When using risk-neutral scenarios, only the average discount value is useful. The distribution implied from risk-neutral scenarios probability is not realistic. Real-world scenarios are used for other purposes. It provides us a picture of possible outcomes. The distribution of the outcome itself, and other risk measures, such as value at risk (VaR), and conditional tail expectation (CTE), are very important tools in capital management and risk-return analysis. Normally, the most common measure is Value at Risk (VaR) which is the value at a certain percentile of a distribution. For example, 95% VaR is the 95 percentile of the loss distribution. Sometimes for distributions that are heavily skewed, VaR may underestimate the tail risk, as it does not consider the magnitude of the loss in the tail. Conditional Tail Expectation (CTE), a.k.a Tail VaR, provides more comprehensive information

---

[40] Jaimungal and Wang (2006) illustrated the analytical tractability when the loss size follows the Gamma distribution. Lin and Chang (2007) illustrated the analytical tractability when the loss size follows the lognormal distribution. Ma and Liu (2004) introduced an analytical scheme to compute the n-fold convolution of exponential-sum distribution functions.

[41] Lanzenauer, "The n-Fold Convolution," 94-98.

about the uncertainty in the tail. Combined with further modeling and understanding of practical accounting nature, such risk analysis would help the investors understand the downside risk and assist the issuers to assess the effectiveness of contingent capital in risk transfer. It can also help quantify the potential benefit of contingent capital under stress scenarios as an integrated part of plan sponsor's capital planning strategy.

Account for Illiquidity

Assets that are less liquid usually have a higher yield to compensate for illiquidity, ceteris paribus. This is often the case for the corporate bond market, where substantial difference is present and the liquidity premium changes through the economic or business cycle. Contingent capital market is still under development and its cash flows are hard to perfectly replicate using existing liquid assets, due to the uncertainty of an embedded conversion option. Realizing the rise of liquidity premium of assets during the recent financial crisis, regulators are considering adding liquidity premium in the valuation of insurance liability as well.[42] Therefore, a liquidity premium is a key factor to consider in setting the price. In the models discussed above, liquidity has not been explicitly included in the pricing framework. The impact of illiquidity can be reflected in the pricing models by adding a liquidity premium to the interest rate. However, even for assets without the conversion option, it is not easy to quantify the liquidity premium. The asset yield in excess of the risk-free rate includes expected credit spread, unexpected credit spread, liquidity premium, cost of conversion option, taxation difference, and residual spread caused by market inefficiency such as information asymmetry.

At the current stage, it is hardly reliable to use the market data of existing contingent capital deals for estimating the liquidity premium, as the data are not sufficient and the terms of the contracts vary greatly. However, the corporate bond market has more data and there are many studies about how to disentangle liquidity premium from the total spread. It is practical to leverage on the liquidity premium of a similar corporate bond with the same credit rating as of contingent capital. Further adjustment can be applied based on the size of the issuance and estimated demand. The following methods may be used for estimating the liquidity premium for corporate bonds.

(1) Market spread - model spread derived from structural models. The model spread is considered as the credit spread.

(2) Market spread - CDS spread with the same maturity and credit rating.

---

[42] CEIOPS, "Task Force Report," 3-12.

(3) Illiquid bond return – liquid bond return with the same cash flows and credit risk.

5.2 Valuation Models and Assumption Setting

Sometimes, pricing is considered as valuation at the issue date. The quantitative models used for pricing and valuation normally are the same. However, due to different purposes and the evolution of the market since the issuance, the parameters used may change greatly. Even for valuation, the assumption used depends on the role of the stakeholder and the purpose of the valuation. This is not something new that is created by contingent capital. It exists for all the complicated financial instruments that do not have a liquid market. This section focuses on the fair valuation of the contingent capital before conversion. Upon conversion, there will be a liability written off or additional equity, whose value may be determined by the market.

The fair value discussed here is the exit price that represents the price to be paid when the asset is transferred, or the price to pay when the liability is transferred. This is the type of fair value that is adopted by IFRS and U.S. GAAP. Several difficulties exist for the contingent capital valuation at current stage.

(1) There is no liquid market for contingent capital right now. Therefore, the exit price is not readily available from the market.

(2) Although there are some ways to dynamically hedge a few risks embedded in contingent capital, the hedging effectiveness is been questioned. The risk exposure could evolve quickly in stressed situation and the dynamic hedging program may not be able to offset the change quickly enough. Other types of risk may not be able to hedge at all, such as systemic risk. Those hard-to-hedge or non-hedgeable risks need to be considered when estimating the fair value.

(3) The exposure to stakeholders' behavior adds an extra layer of complexity. The quantification of its impact is not a trivial task and it is hard to get consensus on the assumption and the conclusion.

There are several approaches that may be used for valuing financial instruments. However, due to the characteristics of contingent capital and its current market, not all of them are appropriate.

(1) Using an up-to-date market price of the contingent capital instrument, which is likely to be unavailable and illiquid in the current market.

(2) Using the market value of asset instruments that can replicate the payments of contingent capital. Due to the embedded conversion option and the stakeholders' behavior, which are

hard to predict, it is difficult to find asset instruments in the market that can replicate the cash flows of contingent capital well in all situations.

(3) Using the discounted value of the expected cash flows with a discount curve that includes the risk premium. This is also known as the actuarial approach. It is extensively used to calculate the value of the insurance business. However, it is not well equipped to value the asset instruments with embedded options.

(a) The life of the contingent capital instruments and its cash flows vary greatly among different financial conditions, due to the conversion option. It is difficult to calculate the appropriate expected cash flows. The cash flows projected under the best estimated assumption may not be a good estimate, as the impact of a conversion option is likely to be neglected or underestimated when it is out of money at the time of valuation.

(b) Since there is not enough active trading in current market, it is difficult to derive the risk premium from the market price. Subjective assumption has to be made and the discount value can hardly be the market consistent value.

(4) An extension of the third approach is taking the average of the discounted values based on real-world stochastic scenarios. The discount factors or the state prices are calibrated to the market price of asset instruments, where possible. In practice, this approach is difficult due to the unyielding number of state prices that need to be determined, especially for a multi-period arbitrage free model. To estimate the market consistent value, risk-neutral scenarios are often used to calculate the non-arbitrage price.

(5) Using the closed form formula where the discount rate is the risk-free rate and model parameters such as equity volatility are calibrated to market price of asset instruments. In this way, it can achieve a certain level of market consistency where liquid market is available. However, it is not easy to account for the impact of stakeholders' behavior.

(6) Taking the average of the discount values of the cash flows which are projected and discounted using risk-neutral stochastic scenarios. The economic scenario models are calibrated to the market price of asset instruments, where possible. In addition, the impact of expected investor behavior can be reflected to a certain extent. For example, near the trigger point, the stock price may be depressed by selling the existing shares or short selling. To account for this, the conversion can be assumed to occur before the stock price reaches the conversion price.

Therefore, the closed-form valuation and stochastic risk neutral valuation are relatively more appropriate for valuing contingent capital instruments.

The models described in Section 5.1 can be used, but with adjusted parameters that reflect the change since the issue date. The major areas that require contemporary assumption are given below.

(1) The economic assumption needs to be updated based on the then-current economic climate. It includes, but is not limited to, interest rates, credit spread, equity value and expected equity return.

(2) For contingent capital that covers non-financial risks such as catastrophe risk, the assumption of the frequency and severity of the risk events also needs to be updated.

(3) The financial condition of the issuer may also change and this could have a material impact on the value of contingent capital with a trigger event at institutional level. The probability of default or conversion is affected by the business strategy and outlook.

(4) Regulatory changes sometimes have an impact on the probability of conversion for contingent capital as well. For example, for the conversion option based on the capital adequacy ratio, changes in the capital rules will directly affect the chance of conversion. Another example is the contingent capital with a discretionary trigger event. Changes in the goal or style of the regulators may also affect their willingness to trigger the conversion.

(5) The volatility assumption of the equity return is also important, And the value of contingent capital is very sensitive to it. In most cases, the volatility term structure is not flat. Even if the volatility curve itself remains unchanged, the originally out-of-money conversion option could become in the money, and if that happens the volatility parameter would need to be updated.

In addition, the occurrence of conversion depends on many factors that cannot be fully controlled or hedged, such as the change in capital rules, business strategy, economic cycle, and business environment. For example, if a CoCo bond has a trigger event based on the regulatory capital adequacy ratio, the timing of the conversion and amount of payment can be affected by a change in the capital rule. The estimated value based on either closed-form valuation or stochastic risk-neutral valuation needs to be reduced to reflect those non-hedgeable risks.

A possible way to determine the amount of the adjustment is to calculate the cost of holding extra capital to cover the risk exposure. The loss can be projected annually, if a tail event related to the non-hedgeable risks happens. The amount of adjustment is the cost of capital rate × present value of estimated annual losses. This is an idea borrowed from the market consistent valuation of insurance liabilities. Cost of residual non-hedgeable risks (CRNHR),[43] a component of market-consistent embedded value (MCEV), uses this approach to estimate the cost due to the exposure to the non-hedgeable risks. If there is a probability associated with the tail event, such as a 1-in-200-year event, or 99.5 percentile of the loss distribution, the capital required in the tail event is in the form of 99.5% value at risk. A more conservative approach to quantify the risk is to use the average loss if it is greater than the 99.5 percentile loss.

The tail event used to quantify non-hedgeable risk, in most cases, is difficult to choose, not to mention assigning a probability. It could be a historical extreme event or a prediction of the future crisis. The associated probability could be based on historical data or predicative models.

5.3 Risk Assessment

5.3.1 Greeks

Greeks are used to measure the sensitivity of financial instruments to key drivers such as equity price, interest rate, volatility, and time. The sensitivity of contingent capital can be estimated by calculating Greeks using valuation models.

(1) Delta (Δ) = $\delta V/\delta S$, where δV is the change in the value of contingent capital and δS is the change in the equity price. The embedded conversion option is highly related to the capital adequacy of the issuer or the whole industry. Therefore, contingent capital instruments, especially those with a fixed conversion price, are sensitive to the issuer's equity price if the trigger is based on an institutional level event, or the industry's equity index if the trigger is based on an industry level event.

(2) Gamma (Γ) = $\delta^2 V/\delta S^2$ explains the convexity of the value with respect to the equity price. The secondary level impact is more material when the equity price moves to a level that the conversion option is close to be exercised.

(3) Vega (ν) = $\delta V/\delta\sigma$, where $\delta\sigma$ is the change in equity volatility. Higher volatility means higher value of the conversion option.

(4) Rho (ρ) = $\delta V/\delta r$, where δr is the change in interest rate. It measures the sensitivity to interest rate. Contingent capital acts like fixed income securities before conversion. The present value of future coupon payments and redemption/conversion value will change if discount rate changes.

---

[43] CFO Forum, "Market Consistent," 5-6.

Given the diversity of contingent capital designs, it is hard to estimate the Greeks using the same formula. One possible way to generalize the estimation of Greeks is to translate the conversion option to a financial derivative with equity as the underlying asset. Valuation models that are analytically tractable can be used to calculate the Greeks using a closed-form formula. Stochastic valuation models can also be used to estimate the Greeks by revaluating the contingent capital instrument under shocked scenarios. In some cases, economic variables are assumed to be interdependent. For example, in an economic recession, a bear equity market is likely to be coupled with low government bond yield. A stochastic model can build the relationship in the scenarios, and it would be a better choice than an analytic model if this kind of interdependency needs to be considered.

5.3.2 Contingent Capital Hedging

Theoretically, the risks of contingent capital can be offset either by static hedging or by dynamic hedging. There have been extensive studies about replicating exotic options with plain vanilla options statically. As long as the contingent capital can be translated into a portfolio of financial instruments, static hedging techniques can be applied.

In a case when there is no market to short the replicating portfolio or the cost is too high, dynamic hedging is another choice. It sets up a hedging portfolio that can offset the sensitivity of the contingent capital. The hedging portfolio needs to be rebalanced as the sensitivity changes, which may be due to a market movement or simply the passage of time.

However, the hedging may not be as effective as expected. The basis risk for some contingent capital instruments may be high. For example, for the CoCo bonds with the trigger event based on the statutory capital adequacy ratio (CAR), the equity price may not be a perfect indicator of the CAR due to the complexity of the capital rules. In addition, the CAR may be reported quarterly while the equity price changes every day. The issuer may go bankrupt before the conversion has a chance to take place. Hedging strategies built for the "translated" contingent capital are vulnerable to basis risk.

In addition, hedging activities may have detrimental market impact. The investors of contingent capital hold a short position on the conversion option. If the conversion price is fixed, short selling of the issuer's stock provides an efficient Delta hedge. This will drag down the equity price further. When the stock price drops, more shares need to be short sold as the delta of the put option increases, which in turn puts more pressure on the equity price.

5.3.3 Earnings Volatility and Capital Adequacy

Investors of contingent capital need to consider the impact on their risk profile. Earning volatility may increase due to the embedded conversion option offered to the issuers. Their available capital might be reduced if the conversion price is greater than the spot price at conversion. The impact will directly show on the financial reports such as income statement and balance sheet. Therefore, it is necessary to project the possible outcomes under different real-world scenarios, either stress scenarios or stochastic scenarios. With a clear understanding about the distribution of the future gain/loss, the portfolio managers can make an informed decision on whether the risks are acceptable or if there is any risk mitigation action to take.

On the other hand, issuers of contingent capital would expect reduced earnings volatility and enhanced capital position upon conversion.

6. CASE STUDY

In this section, we will go through the pricing, valuation, and risk analysis of a sample contingent capital instrument. It is hoped that the reader will gain some perspectives of the fundamental quantitative works required for analyzing contingent capital. Given the diversity in the features of contingent capital, the methods used in the case study may not be enough or appropriate for other types of the contingent capital but the principle will not deviate too much. Although there is some calibration involved in the case study, it is insufficient to ensure a reasonable and marketable price or value. The main purpose of the case study is to illustrate the model, the process and the Excel tool built with it. More detailed market research is required to come with appropriate model parameters.

The details of the contingent capital instrument example are given below.

Facts about CoCo Bond XYZ	
Issuer	ABC Insurance Company
Face Amount	$10,000,000
Trigger Event	NAIC RBC Ratio <=150%
Conversion Price (CP)	$40 per share
Term of Contract (T)	10 years
Current Stock Price (S_0)	$45 per share
Current RBC Ratio (RBC_0)	300%

Credit Rating	S&P BBB+
BBB+ rated junior subordinated bond yield	7.2%
Dividend Yield (d)	0%
Credit Default Swap Curve for subordinated bond	Term / Rate (bps) 1 / 124.9 2 / 198.9 3 / 262.5 4 / 300.6 5 / 311.8 7 / 313.9 10 / 305.0
Economic Assumption	
Risk Free Interest Rate[44] (r)	3.0%
Equity Volatility (σ)	45%
Recovery Rate for Junior Subordinated Bonds	40%

6.1 Pricing

The task is to determine an appropriate and marketable coupon rate for CoCo bond XYZ.

A plain vanilla junior subordinated bond with the same credit rating as the CoCo bond has a yield of 7.2%. Since the conversion price is well above the likely stock price at the conversion, the exercise of the conversion option means a loss to the CoCo bond holders. Therefore, the required yield of CoCo bond XYZ needs to be higher than 7.2% in this example.

[44] In the case study, liquidity premium is assumed to be accounted for in the risk-free interest rate, where appropriate. The determination of liquidity premium is discussed in Section 5 and is not illustrated here.

CoCo bond XYZ has a trigger event linked to the NAIC RBC ratio. The distribution of future RBC ratio is difficult to model due to the complexity of the liability portfolio of insurance companies, unpredictable changes in capital rules, and the uncertainty of management actions. On the other hand, quantitative financial models explicitly project future stock price, interest rate, and credit spreads. To rely on those models, the relationship between the RBC ratio and those modeled economic variables need to be figured out. Two likely explanatory variables are stock price and credit default swap (CDS) rates. Assume that for ABC Insurance Company, the stock price is expected to be \$15 at conversion based on the historical data of RBC ratio and stock price, an expected capital rule change, and the next 5-year risk budgeting plan.

The Spiegeleer and Schoutens (2011) credit derivative approach

The Spiegeleer and Schoutens (2011) credit derivative approach can be used to get a rough estimate of the CoCo bond yield. The probability of triggering during the term of the contract can be calculated using the valuation formula for down-and-in cash (at expiry)-or-nothing binary option[45] without discounting the payoff back to the valuation date.

$$p = \Phi(d_1) + \left(\frac{B}{S}\right)^{2\alpha} \Phi(d_2)$$

$$d_1 = \frac{\log\left(B/S\right) - \left(r - d - \frac{\sigma^2}{2}\right)\cdot T}{\sigma\sqrt{T}} \qquad d_2 = \frac{\log\left(B/S\right) + \left(r - d - \frac{\sigma^2}{2}\right)\cdot T}{\sigma\sqrt{T}}$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}\, dy \qquad \alpha = \frac{r-d}{\sigma^2} - \frac{1}{2}$$

Parameters	B (stock price at conversion)	S	T	R	d	σ
Value	15	45	10	3%	0%	45%

The probability of triggering during the life of the CoCo bond is estimated to be 61.3% based on the parameters listed in the table above. This indicates an intensity of 0.095 for the triggering based on the formulae given below. The recovery rate at conversion can be calculated as the ratio of the

[45] Rubinstein, Mark and Eric Reiner, "Unscrambling," 75-83.

stock price at conversion and the conversion price. It is 37.5% (15/40) in the example. $\lambda = -\frac{\log(1-p)}{T}$

Credit spread of CoCo Bond XYZ = λ × (1-Recovery Rate) = 5.9%

Total yield of CoCo Bond XYZ = credit spread + risk free rate = 8.9%

The Spiegeleer and Schoutens (2011) equity derivative approach

With an explicit consideration of the time of the conversion and the stop of coupon payment upon conversion, equity derivative approach determines the price of the CoCo bond as below.

(1) Plain Vanilla Bond Price with risk-free discounting

\+ (2) Knock-In Forwards between spot price and conversion price[46]

\- (3) Down-and-in cash-or-nothing binary option on the coupon payments

The value of knock-in forwards and the binary option on the coupon payments can be calculated based on the well-established pricing formula of binary options. Using the pricing formula given by Spiegeleer and Schoutens (2011),[47] an annual coupon rate of 9.3% will make the price of CoCo bond XYZ equal to its face amount. This is different from the total yield of 8.9% derived using the credit derivative approach. As mentioned in Section 5, credit derivative approach neglects the impact of coupon payments and might generate an unrealistically low bond yield, especially when the recovery rate is high.

Garcia and Pede (2011) analytical first passage time approach

The analytical first-passage time approach model enhances the Merton model by introducing a time-dependent barrier for default and a non-flat volatility term structure. Details about the model are provided in Appendix B.

$V_t = S_t + \hat{H}(t)$.

Firm-value process: $dV_t = rV_t dt + \sigma(t) V_t dW_t^Q$

Barrier: $\hat{H}(t) = He^{rt - B\int_0^t \sigma^2(s)}$

Equity: $S_t = P_t \mathrm{E}_t\left[\frac{\left(V_T - \hat{H}(T)\right)^+ 1_{\{\tau > T\}}}{P_T}\right]$

---

[46] The knock-in forward will be effective if the trigger event for the CoCo bond happens.

[47] Spiegeleer and Schoutens, "Pricing Contingent," 2011, 24. It is also listed in APPENDIX B.

RBC ratio: $$RBC_t = g\left(V_t / \hat{H}_t\right) + \varepsilon_t$$

Model parameters B, H, and σ (t) are calibrated to credit default swap (CDS) spread, equity value, and capital adequacy ratio.

Parameters	B	H	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7
Value	0.03	0.7	15.0%	11.0%	12.0%	12.4%	11.6%	11.3%	11.2%

The CoCo bond price can be calculated using the following process.

(1) Simulate the firm value and barrier.

(2) Conversion time τ_c is simulated based on the value of V_t/H_t compared to a threshold translated from the RBC trigger level.

(3) If there is no conversion before bond maturity, the value is the same as the value of the plain vanilla bond with risk-free discount rate. If there is a conversion, it is calculated as the value of paid coupons and the value after conversion.

(4) Take the average of the bond value across all scenarios.

With the following model setup, a CoCo bond with an annual coupon rate of 8.7% will sell at par.

Parameters	RBC Ratio Report Frequency	Bond Maturity	# of Scenario	Conversion Price	Value at Start	V/H threshold
Value	annual	10	1000	0.27	Firm: 1 Barrier: 0.7 Equity: 0.3	120%

More frequent capital adequacy reporting will lead to a lower CoCo bond price.

RBC Ratio Reporting Frequency	Annual	Semi-annual	Quarterly
Value	1.00	0.98	0.96

95% Confidence Interval	0.97 ~ 1.03	0.95 ~ 1.01	0.93 ~ 0.99

If there is a strong belief that the death spiral will happen when the company approaches the trigger level, the model can be adjusted by increasing the asset volatility and assume there is only downward movement when V/H is close to the threshold. The impact of death spiral on the bond value is quite material based on the illustration given below.

Investor Behavior	No	Short Selling near Trigger Level*
Value	1.00	0.85
95% Confidence Interval	0.97 ~ 1.04	0.82 ~ 0.88

* Downward movement and twice the calibrated volatility are assumed when V/H is below 125%.

If there is an expectation of some ad hoc changes before the CoCo bond matures, jumps can be added to both the firm value and barrier. The goal of adding the jump component is to incorporate the expectation of more stringent capital requirement in the near future and the management actions in reducing the resulting cost by adjusting business strategy and mitigating risks.

Firm-value process: $V_t = V_0 \exp\left(rdt + \sigma(t) W_t^Q + \sum_{i=1}^{N_i} Y_i \right)$

Barrier: $\hat{H}(t) = He^{rt - B\int_0^t \sigma^2(s) + \sum_{i=1}^{N_i} Z_i}$

N determines the number of jumps and it follows the Poisson process with parameter λ

Y determines the shock size due to management actions. In this example, it is assumed to be positive but less than Z.

Z determines the shock size due to capital rule changes. In this example, it is assumed to be positive to account for more stringent capital requirement in the future.

For the sake of simplicity, the following parameters are used in the example. A compound Poisson process is simulated to determins both the number of jumps before bond maturity and the arrival time of jumps.

Parameters	λ	Fixed size of Y	Fixed size of Z	Equity Value at Start
Value	0.5 (expected one time in two years)	0.02	0.05	0.30

The value of CoCo bond is sensitive to those overall negative jumps.

Jumps	No	Compound Poisson Process with Fixed Shock Size*
Value	1.00	0.85
95% Confidence Interval	0.97 ~ 1.04	0.82 ~ 0.89

* One in two years with an overall impact of 5.2% drops in equity value per time.

The Duffie and Singleton (1999) Approach with Equity Price State Variable

Duffie and Singleton (1999) proposed a new approach to model financial instruments that are subject to default risk. The default-adjusted short-rate process was introduced which explicitly consider the default hazard rate and loss ratio. The beauty of this model framework is the capability of having a state dependent default hazard rate and loss-ratio process. For a CoCo bond, an ideal candidate of the state variable is the stock price, as the exercise of the conversion option and the value of payoff are correlated with the stock performance. The state process could also be a jump diffusion process which is flexible enough to model ad hoc changes. The following model set up is used in the case study.

Default-adjusted Discount Rate: $R(t) = r(t) + Lh(S,t)$

Stock Price: $S_t = S_0 \exp\left(r(t)dt + \sigma_s dW + \sum_{i=1}^{N_i} Y_i \right)$

Value of Convertible Security: $V_t = \mathrm{E}_t^Q \left[e^{-\int_t^T R(u)du} X + \int_t^T e^{-\int_t^T R(u)du} dC_s \right]$

Conversion Hazard Rate: $h(S,t) = \theta + \frac{\rho}{S_t}$

Loss Ratio at Conversion: $L = (1 - K / CP)$

<u>Notations:</u>

X: Redemption Value

C_t: Coupon Payment Process

CP: Conversion Price

$\sum_{i=1}^{N_i} Y_i$: Jump Component that follows Compound Poisson Process with negative shock size.

K: Translated threshold for stock price at or below which the conversion option will be exercised.

The volatility parameter of equity price process can be calibrated using equity option market value. The jump component can be used to model expected future discontinuous changes. A translated threshold for equity price is used to approximate the trigger event. A fixed recovery rate is assumed as the translated threshold for equity price divided by the conversion price. This implicitly assumes that the exercise of the conversion option is continuous. In reality, stock price could drop well below the translated threshold before the occurrence of the trigger event. It can be compensated for by adjusting up the hazard rate for conversion. However, the key challenge of using this method is the calibration of the conversion hazard rate function. Due to the lack of liquidity in contingent capital market, it might be difficult to have something market consistent. A possible way of estimation is given below.

Step 1: Calibrate the parameters (θ and ρ) to the price of a plain vanilla bond without the conversion option;

Step 2: Estimate the probability of conversion before bond maturity, based on the translated threshold for equity price;

Step 3: Adjust ρ to be the estimated ρ in step 1 × probability of conversion/probability of default.

With the following inputs and parameters, the CoCo bond is priced at par in this example.

K (estimated threshold for stock price)	Coupon Rate	Plain Vanilla Bond Yield	Hazard Rate Function		S	T	R	σ	Jump□	
			θ	ρ□					λ	Size
15	8.7%	7.2%	0.035	1.46	45	10	3%	45%	0.5	0.05

Notes:

1. A ρ of 0.92 generates a model value of the plain vanilla bond price equal to its market value. The stock price process is used to approximate the probability of conversion and the probability of default by assuming that a stock price of $15 or less leads to a conversion and a stock price of $10 or less leads to a default.
2. One in two years with an overall impact of 5.2% drop in equity value per time.

6.2 Valuation

The valuation process is quite similar to pricing except that the economic environment and financial condition could be much different from those at the issue date. Since there are some non-hedgeable risks embedded in contingent capital, its market is not a complete market. Therefore, there is a need to deduct the cost of residual non-hedgeable risks (CRNHRs) from the value calculated using the pricing model. The way of calculating a market-consistent embedded value of insurance products can be borrowed to estimate the cost.

A common method used to estimate CRNHR is the Cost of Capital (CoC) approach.

$$CRNHR = \sum_{t=0}^{T-1} REC_t \times CoC \times v_{t+1} \times p_t$$

REC : Required Economic Capital for Non Hedgeable Risks

CoC : Cost of Capital

v_t : Discount Factor at time t

p_t : Survival Probability

REC_0 = shocked CoCo bond value under stress scenario - current CoCo bond value

REC_t can be estimated as $\alpha \times \text{Risk Driver}_t$

$\alpha = REC_0 / \text{Risk Driver}_0$

Continue with the pricing example of using the analytical first-passage time approach with annual RBC reporting frequency. The non-hedgeable risk to consider is a more stringent capital rule. The following jump component is used to represent the stress scenario.

Parameters	λ	Fixed size of Y	Fixed size of Z	Equity Value at Start
Value	0.5 (expected one time in two years)	0.06	0.15	0.30

* One in two years with an overall impact of 17.1% drops in equity value per time.

Under the stressed scenario, the model value of a CoCo bond becomes 0.657. The risk driver is set to be the price of a plain vanilla bond without the conversion option and assuming no default risk.

α	CoC Rate	CRNHR	Adjusted CoCo Bond Price
25.9%	4%	0.09	0.91

When more than one non-hedgeable risk is considered, correlation among the risks needs to be quantified to reduce the aggregated required economic capital.

6.3 Risk Assessment

Greeks

Greeks are used to illustrate the sensitivity of the CoCo bond value to economic variables or model assumptions. Using Garcia and Pede's (2011) analytical first-passage time approach, the estimated Greeks are given in the table below. A negative Gamma means that the second order impact of a drop in equity value on CoCo bond price is also negative. This is expected, as an equity price decrease will not only increase the probability of conversion but also push forward the timing of the conversion.

Parameter				Bond Price			Greeks	Value
Scenario	Baseline	Up	Down	Baseline	Up	Down	**Δ: Delta**	**2.6**
Equity Value	0.3	0.31	0.29	1.00	1.02	0.97	**Γ: Gamma**	**-31.1**
Interest Rate	3%	0.03	0.03	1.00	1.00	1.01	**ρ: Rho**	**-6.2**
Volatility		+ 1%	- 1%	1.00	0.95	1.05	**ν: Vega**	**-4.8**

Stochastic Analysis

For the investors, it is also useful to take a look at the distribution of the bond value and the worst possible outcome. On the other hand, the issuer would be interested to know how much capital relief it could gain from the conversion option and how it can offset the negative earnings in the stressed situation.

Stochastic analysis with real-world scenarios can be used to look into the tail risk of investing in a CoCo bond and the benefit of issuing a CoCo bond at the tail event. It is also an appropriate framework for incorporating an expectation different from the market. In the following example, under some arbitrary real-world economic assumptions, the stock price, conversion time, and the loss of the investor at conversion are simulated and the outcomes at the tail are summarized. One thousand scenarios are used for illustration.

Model Setup

Risk Free Rate: $dr = \{\theta(t) - \alpha\, r(t)\}dt + \sigma_r dW_r$ (One Factor Hull White Model)

Stock Price: $S_t = S_0 \exp\left(\mu(t)dt + \sigma_s dW_s + \sum_{i=1}^{N_i} Y_i \right)$

Expected Equity Return: $\mu(S,t) = r(t) + risk\ \ premium$

Correlation of Diffusion Processes: $\rho = corr(dW_r, dW_s)$

Conversion Time: $\tau_c = \inf\{t \geq 0\ s.t.\ S_t \leq K\}$

Loss Ratio at Conversion: $L(S,t) = 1 - Min(K, S_{\tau_c})/CP$

<u>Notations:</u>

CP: Conversion Price

$\sum_{i=1}^{N_i} Y_i$: Jump Component that follows Compound Poisson Process with negative shock size.

K: Translated threshold for stock price at or below which the conversion option will be exercised.

The following table lists major parameters used in the example.

K (estimated threshold for stock price)	Coupon Rate	Plain Vanilla Bond Yield	One Factor Hull White		S	T	R	ρ	σ	Jump☐	
			α	σ						λ	Size
15	8.7%	7.2%	0.9	1.5%	45	10	3%	5%	45%	0.5	0.05

Notes: One in two years with an overall impact of 5.2% drop in equity value per time.

Value of tail risk measures are given below. A higher loss ratio for the investors means a greater reduction of debt obligation for the issuer.

Variables / *Risk Measures*	Stock Price at Conversion / *Left Tail*	Conversion Time / *Left Tail*	Loss Ratio at Conversion / *Right Tail*
95% VaR	9.46	3.00	76%
95% CTE	8.14	2.30	80%
99% VaR	7.22	2.00	82%
99% CTE	6.53	1.60	84%
99.5% VaR	6.92	2.00	83%
99.5% CTE	5.94	1.20	85%

Figures 1-3 illustrate the probability density functions of conversion time, stock price at conversion, and loss ratio, given that the conversion happened.

Figure 1: Histogram of Conversion Time Given It Happens

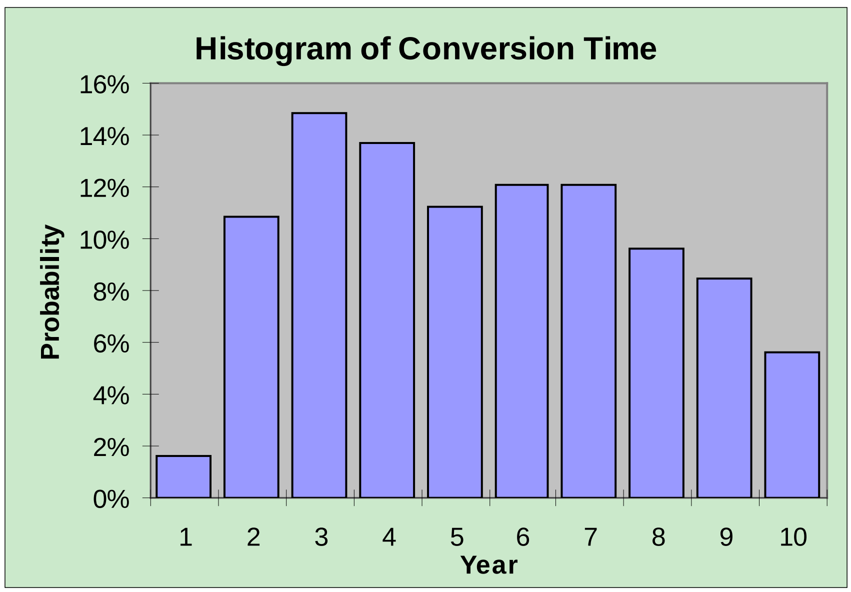


Figure 2: Histogram of Stock Price at Conversion

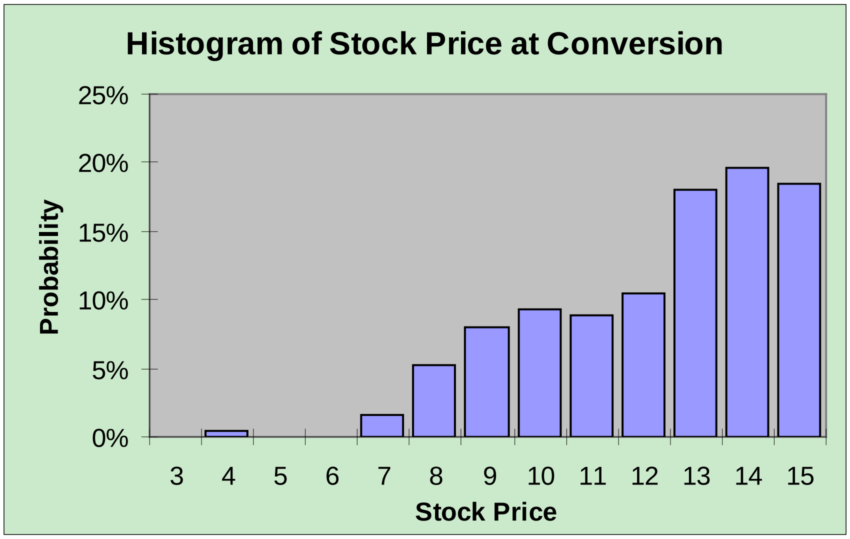
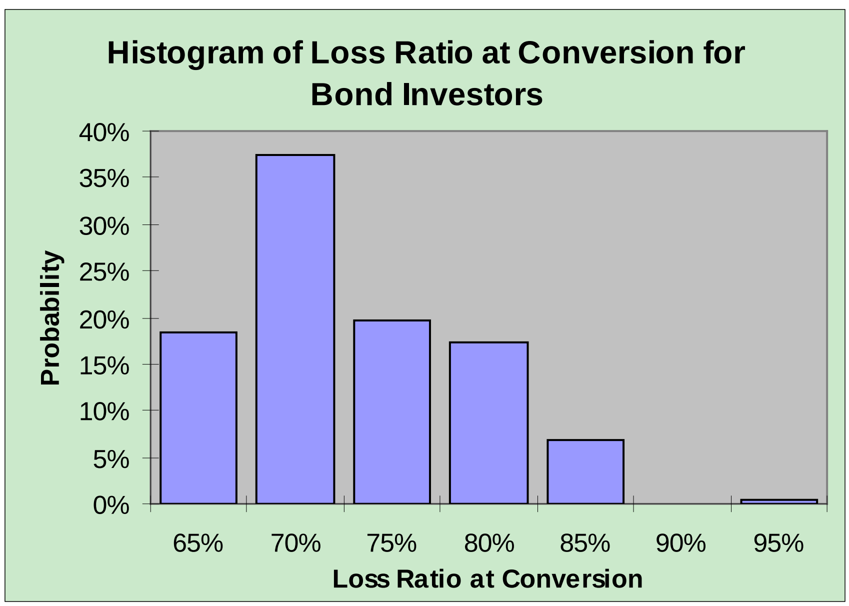


Figure 3: Histogram of Investor's Loss Ratio at Conversion



Based on the purpose of the risk analysis, it can be extended to quantify the impact on earnings volatility and capital at risk, based on a specified financial reporting framework and capital rules.

7. CONCLUSION

Contingent capital is considered as a promising candidate for improving the risk tolerance of the financial industry and reducing the cost of the financial crisis paid by the taxpayers. Compared to subordinated debt instruments, contingent capital increases the capability of absorbing loss. Compared to equity, contingent capital has a lower cost of capital before conversion. Despite the doubts about its success, it is welcomed by the regulators and there have been many proposals of the appropriate design of contingent capital instruments.

However, there is still a long way to go before contingent capital can be widely accepted and utilized.

(1) The trigger event has so many possibilities that choosing an appropriate design is not an easy task. A small change of the feature may have a material impact on its effectiveness of reducing the chance of default. There remains a lot to discover and test in the market.

(2) Closely related to the trigger event, the behavior of the stakeholders needs more analysis. They include both rational behaviors, and irrational behaviors such as panic. Some behaviors may drag the issuer down further near conversion instead of helping as intended. They need to be fully understood and the potential impact needs to be quantified.

(3) The complexity and uncertainty of contingent capital make it difficult for pricing, valuation, and risk assessment. Although there are some models for analyzing contingent capital, they are highly data driven. Those garbage-in garbage-out models will not be very useful before a liquid market emerges for contingent capital. How to set a fair price is more of an art than a math problem.

Hopefully after those issues are solved, contingent capital will be instrumental in reducing the systemic risk of the industry and default risk of financial institutions without incurring too much additional cost of capital for investors.

Acknowledgment

The author would like to thank VFIC members Chris Gross, Edward Yao, Philip Kane, and Rasa Mckean (Chair) for their guidance, review, comments, and full support of this research. Without their insightful input, this paper would not have its current level of sophistication and relevance to potential readers. Special thanks are given to Professor Shaun Wang and Mr. Han Chen for their academic review of the report and thought-provoking advice on the appropriate pricing methods. The author is grateful for the opportunity provided by the Committee on Valuation, Finance, and Investments (VFIC) of Casualty Actuarial Society (CAS). The authors also would like to thank David Core for his coordination on this project.

Supplementary Material

A spreadsheet is built to illustrate the pricing, valuation, and risk analysis for contingent capital. It is used intensively in the case study.

APPENDIX A. QUICK GUIDE FOR CONTINGENT CAPITAL QA TOOL

CONTINGENT CAPITAL QA TOOL is a spreadsheet model built to illustrate the pricing, valuation, and risk analysis for contingent capital. It is capable of pricing/valuing certain types of contingent capital instruments using closed form solution and stochastic approach. Model calibration and risk analysis are also included. It could serve as a good education material to understand contingent capital and risk quantification. All the quantitative results used in the case study are generated using this tool.

In order to use this tool properly, the user needs to enable the macros after the spreadsheet is opened. Most of the calculation functions are built using VBA. In addition, the user needs to accept the disclaimer statements before using the tool. The spreadsheet has most of its input cells green colored and output cells blue colored. Tab "ReadMe" provides descriptions of the functionality, output, and new functions built with VBA.

The following models have been built in the tool.

(1) Spiegeleer et al. (2011) Credit Derivative Approach[47]: Tab "S&S Credit Approach";
(2) Spiegeleer et al. (2011) Equity Derivative Approach[47]: Tab "S&S Equity Approach";
(3) Garcia and Pede (2011) Analytical First Passage Time Approach[27]: Tab "AFPT";
(4) Duffie and Singleton (1999) Approach[15] with equity price as the state variable[50]: Tab "ADS";
(5) Risk Analysis such as VaR and Greeks: Tab "Risk Analysis".

APPENDIX B. MORE FORMULAS USED IN THE CASE STUDY

Spiegeleer and Schoutens (2011) equity derivative approach[48]

CoCo Bond Price

= (1) Plain Vanilla Bond Price with risk free discounting

\+ (2) Knock-In Forwards between spot price and conversion price

\- (3) Down-and-in cash (at expiry)-or-nothing binary option on the coupon payments

$$(1): \quad price\ N \cdot e^{-rT} + \sum_{i=1}^{m} c_i \cdot e^{-rt_i}$$

$$(2): \quad payoff = {}^{N}\!/_{CP} \cdot \left(S(\tau) - CP\right) \ at\ time\ \tau\ if\ \ for\ the\ \ first\ time\ \tau < T_{i,}, S(\tau) < B$$
$$= 0\ if\ \ for\ all\ \tau < t_{i,}, S(\tau) > B$$

It is approximated as the payoff of

(a) down-and-in asset-or-nothing call

[48] Spiegeleer and Schoutens, "Pricing Contingent," 2011, 24.

- (b) down-and-in asset-or-nothing put
- (c) down-and-in asset(at expiry)-or-nothing

The only difference is in the payoff time. For a CoCo bond, it happens at the time of conversion. For (a), (b), and (c), the payoff happens at time T. When the issuer is in a financial distress, the chance of the conversion is high and the conversion is expected to be early if it happens. In this case, the price of using (a)-(b)-(c) may not be a good estimator and may need to be adjusted to avoid an underestimation of CoCo bond price.

$$(2):\ price\quad N/CP\left\{\begin{array}{l} S\cdot e^{-qt}\cdot\left(\frac{B}{S}\right)^{2\lambda}\Phi(y_1)+S\cdot e^{-qt}\cdot\Phi(-x_1) \\ -K\cdot e^{-rT}\left[\Phi\left(-x_1+\sigma\sqrt{T}\right)+\left(\frac{B}{S}\right)^{2\lambda-2}\Phi\left(y_1-\sigma\sqrt{T}\right)\right]\end{array}\right\}$$

$$\begin{aligned}(3):\ payoff &= c_i\ if\ \ for\ all\ \tau < t_{i,}, S(\tau) > B \\ &= 0\ if\ \ for\ some\ \tau < t_{i,}, S(\tau) < B\end{aligned}$$

$$price\quad \sum_{i=1}^{m} c_i\cdot e^{-rt_i}\cdot\left[\Phi\left(-x_{1i}+\sigma\sqrt{t_i}\right)+\left(\frac{B}{S}\right)^{2\lambda-2}\Phi\left(y_{1i}-\sigma\sqrt{t_i}\right)\right]$$

B : Stock price at conversion CP : Conversion price
T : Term of CoCo bond t_i : ith coupon payment time
c_i : ith coupon payment amount N : Face amount
m : # of future coupon payments till time T

$$x_1 = \frac{\log\left(S/B\right)+\lambda T}{\sigma\sqrt{T}} \qquad y_2 = \frac{\log\left(B/S\right)+\lambda T}{\sigma\sqrt{T}}$$

$$x_{1i} = \frac{\log\left(S/B\right)+\lambda t_i}{\sigma\sqrt{t_i}} \qquad y_{2i} = \frac{\log\left(B/S\right)+\lambda t_i}{\sigma\sqrt{t_i}}$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}dy \qquad \lambda = \frac{r-d+\sigma^2/2}{\sigma^2}$$

Garcia and Pede (2011) analytical first passage time approach

The following model and calibration method are applied in the case study. They are based on Garcia and Pede (2011), either identical or with a slight difference.

Model Setup

Firm-value process: $dV_t = rV_t dt + \sigma(t) V_t dW_t^Q$

Barrier: $\hat{H}(t) = He^{rt - B\int_0^t \sigma^2(s)}$

Equity: $S_t = P_t \mathrm{E}_t \left[\frac{\left(V_T - \hat{H}(T)\right)^+ 1_{\{\tau > T\}}}{P_T} \right]$

$V_t = S_t + \hat{H}(t)$

Survival probability: $\mathrm{P}(\tau > T) = \Phi(d_1) - \left(\frac{H}{V_0}\right)^{2B-1} \Phi(d_2)$

CDS model spread: $$S^{T_0,T_N}(t) = \frac{(1-R)\sum_i \frac{P_t}{P_{T_i}} \left(P(\tau > T_{i-1}) - P(\tau > T_i)\right)}{\sum_i \frac{P_t}{P_{T_i}} (T_i - T_{i-1}) \left(P(\tau > T_i) + \frac{1}{2}\left(P(\tau > T_{i-1}) - P(\tau > T_i)\right)\right)}$$

Notations:

$$d_1 = \frac{\log \frac{V_0}{H} + \frac{2B-1}{2}\int_0^T \sigma(s)^2}{\left(\int_0^T \sigma(s)^2\right)^{1/2}} \qquad d_2 = d_1 - \frac{2\log\frac{V_0}{H}}{\left(\int_0^T \sigma(s)^2\right)^{1/2}}$$

$\tau = \inf\left\{t \geq 0 \ s.t. V_t \leq \hat{H}(t)\right\}$ $\qquad P_t = e^{-rt}$ zero coupon bond price with term t

It is also assumed that NAIC RBC Ratio can be estimated based on the value of V_t and H_t.

$RBC_t = f\left(V_t / \hat{H}_t\right) + \varepsilon_t$, where *f* is a monotonically increasing function.

Equity price $E_t = f(V_t, t)$ follows the following partial differential equation and boundary conditions.

$$\begin{aligned}
&\partial_t f = -\frac{1}{2}\sigma(t)^2 x^2 \partial_{xx} f - rx\partial_x f + rf, t \in (0,T) \text{ and } x \in R_+ \\
&f(x,T) = \left(x - \hat{H}(T)\right)^+, x \in R_+ \\
&f(x,t) = 0, 0 \leq x \leq \hat{H}(t), t \in (0,T) \\
&x = V_t
\end{aligned}$$

To transform the boundary condition $f(x,t) = 0, 0 \leq x \leq \hat{H}(t), t \in (0,T)$ to a fixed one, let $E_t = f(V_t, t) = f^*\left(V_t / \hat{H}(t), t\right) = f^*\left(x^*, t\right)$. It then follows the following partial differential equation and boundary conditions.

$$\partial_t f^* = -\frac{1}{2}\sigma(t)^2 x^{*2}\partial_{x^*x^*} f^* - \left(rx^* - x^* \frac{\hat{H}'}{\hat{H}}\right)\partial_{x^*} f^* + rf^*, t \in (0,T)\ and\ x^* \in R_+$$
$$f^*(x^*,T) = \hat{H}(T)(x-1)^+, x^* \in R_+$$
$$f^*(x^*,t) = 0, 0 \le x^* \le 1, t \in (0,T)$$

Calibration Process

Parameters: $H, B, \sigma(i), i = 1 \text{ to M}$

Targets: CDS spreads, equity price, and NAIC RBC ratio.

$$Step\ 1: RBC_0 = g\left(V_0 / \hat{H}_0\right) \Rightarrow H$$
$$Step\ 2: Minimize\ D = \sum_{i=1}^{M}(\text{CDS model spread - CDS market spread})^2 \Rightarrow \sigma(i), i = 1 \text{ to M}$$
$$Step\ 3: E_t \Rightarrow B$$

Return to Step 2 if D is greater than the tolerance level of error

This method does not guarantee a global minimum being found and different guess of initial values need to be tried. Adjusted Levenberg-Marquardt algorithm is used for Step 2.

CoCo Bond Price

The CoCo bond price can be calculated by simulating the conversion time τ_c first, calculating the value using the formulae given below, and then taking the average across the scenarios.

$$CBP(t,T) = 1_{\{\tau_c > T\}} BP(t,T) + 1_{\{\tau_c < T\}}\left(CP(t,\tau_c) + \frac{E_{\tau_c}}{CP} P(t,\tau_c)\right)$$

Where

CBP(t,T): CoCo bond price at time t with bond maturity at time T.

BP(t,T): Risk free bond price at time t with bond maturity at time T.

CP(t, τ_c): The value of coupon payments at time t with payments until time τ_c.

P(t,T): Zero coupon bond price at time with bond maturity at time T.

CP: Conversion price.

E_{τ_c} : Stock price at conversion.

τ_c: Conversion time.

Assume the RBC ratio threshold is RBC*. The corresponding ratio of firm value to the barrier is $g^{-1}(RBC^*)$ at or below which the conversion is triggered. g^{-1} is the inverse function of *g*. Under each scenario, τ_c is determined as the first time $V_t \le g^{-1}(RBC^*)$ or never.

One-factor Hull White Model[49]

$$dr = \left(\theta(t) - \alpha r\right)dt + \sigma dz$$

The mean of the short rate r reverts to $\theta(t)/a$ at rate a.

The following stochastic process for R(t) with annual step is implemented to generate stochastic short rates for stochastic risk analysis.:

$$R(t+1) = bR(t) + \theta_t + \sigma_d \varepsilon_t$$

Auto Correlation: b=e$^{-\alpha}$

Discrete Volatility Parameter: $\sigma_d = \dfrac{\sigma(1-e^{-\alpha})}{\alpha}\sqrt{\dfrac{1-e^{-2\alpha}}{2\alpha}}$

Random Shock: $\varepsilon_t \sim$ N(0,1)

$R(t+1)$ follows normal distribution $N\left(\theta_t + bR(t)\,;\,\sigma_d^2\right)$ given a known $R(t)$

$$R(T) = b^{T-t}R(t) + \sum_{k=t}^{T-1} b^{T-1-k}\left(\theta_k + \sigma_d \varepsilon_k\right)$$

Probability distribution of R(T) at time t is a normal distribution.

Mean: $R(T) = b^{T-t}R(t) + \sum_{k=t}^{T-1} b^{T-1-k}\theta_k$

Variance: $\mathrm{Var}\left(\sum_{k=t}^{T-1} b^{T-1-k}\sigma_d \varepsilon_k\right) = \sigma_d^2\left(\dfrac{1-b^{2(T-t)}}{1-b^2}\right)$

Under risk-neutral valuation, the bond price is calculated as

$$P(t,T) = \mathbf{E}_t^*\left[\exp\left\{-\sum_{k=t}^{T-1} R(k)\right\}\right]$$

$$\sum_{k=t}^{T-1} R(k) = \sum_{k=t}^{T-1}\left(b^{k-t}R(t) + \sum_{l=t}^{k-1} b^{k-1-l}\left(\theta_l^* + \sigma_d \varepsilon_l\right)\right)$$
$$= \left(\frac{1-b^{T-t}}{1-b}\right)R(t) + \sum_{l=t}^{T-2}\left(\frac{1-b^{T-l-1}}{1-b}\right)\left(\theta_l^* + \sigma_d \varepsilon_l\right)$$

Bond price can be written as follows.

---

[49] Hull, J. and A. White: "One-Factor Interest Rate," 235-254. This section describes the details of One-factor Hull White interest rate model that are used for the case study.

$$P(t,T) = \mathbf{E}_t^*\left[\exp\left\{-\left(\frac{1-b^{T-t}}{1-b}\right)R(t) - \sum_{l=t}^{T-2}\left(\frac{1-b^{T-l-1}}{1-b}\right)\left(\theta_l^* + \sigma_d \varepsilon_l\right)\right\}\right]$$

$$= \exp\{-B(T-t)R(t) + A(t,T)\}$$

$$B(S) = \left(\frac{1-b^S}{1-b}\right)$$

$$A(t,T) = -\sum_{l=t}^{T-2}\left(\frac{1-b^{T-l-1}}{1-b}\right)\theta_l^* + \tfrac{1}{2}\sigma_d^2 C(T-t-1)$$

$$C(S) = \frac{1}{(1-b)^2}\left(S - 2b\left(\frac{1-b^S}{1-b}\right) + b^2\left(\frac{1-b^{2S}}{1-b^2}\right)\right)$$

Term structure of interest rates (zero's) at time t is given by solving P(t,T) = exp{-Z(t,T)(T-t)}:

$$Z(t,T) = \frac{-\ln P(t,T)}{T-t}$$

$$= \left(\frac{B(T-t)}{T-t}\right)R(t) - \left(\frac{A(t,T)}{T-t}\right)$$

Model Calibration

Interest rate model parameters $\theta_t, t = 0,1,2,...$ need to be calibrated to the initial yield curve. It can be achieved by solving the following function. Z(0,T) is the T-year zero rate at time zero.

$$-Z(0,T)T = -\sum_{l=0}^{T-2}\left(\frac{1-b^{T-l-1}}{1-b}\right)\theta_l^* + \tfrac{1}{2}\sigma_d^2 C(T-1) - B(T)R(0)$$

This equation can be solved iteratively. For T=2,

$$\theta_0^* = Z(0,2)2 + \tfrac{1}{2}\sigma_d^2 C(1) - B(2)R(0)$$

The values for θ_1^*, θ_2^*, etc. can be found for T=3,4,…

$$\theta_{T-2}^* = Z(0,T)T - \sum_{l=0}^{T-3}\left(\frac{1-b^{T-l-1}}{1-b}\right)\theta_l^* + \tfrac{1}{2}\sigma_d^2 C(T-1) - B(T)R(0)$$

Market instruments such as cap/floor and swaption can be used to calibrate the volatility parameters. Given the analytical tractability for the one-factor Hull White model, it is relatively easy to calibrate the parameters using either closed form formula or trinomial tree model[50].

8. REFERENCES

[1.] Acharya, Viral V, Thomas Cooley, Matthew Richardson, and Ingo Walter, "REAL TIME SOLUTIONS for

[50] Hull, J. and A. White, "Using Hull-White," 26-36.

FINANCIAL REFORM," An NYU Stern Working Group on Financial Reform, 2009, 42-43.
http://govtpolicyrecs.stern.nyu.edu/docs/whitepapers_ebook_full.pdf

[2.] Admati, Anat R., Peter M. DeMarzo, Martin F. Hellwig, and Paul Pfleiderer, "Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Regulation: Why Bank Equity is Not Expensive," 2011, 59-62.
http://www.coll.mpg.de/pdf_dat/2010_42online.pdf

[3.] Basel Committee on Banking Supervision, "Global systemically important banks: assessment methodology and the additional loss absorbency requirement, Rules text," 2011, 18-21, 26.
http://www.bis.org/publ/bcbs207.pdf

[4.] Bolton, Patrick and Frederic Samama, "Capital Access Bonds: Contingent Capital with an Option to Convert," 2011, 3-47.
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1960505

[5.] Calomiris, Charles W. and Richard J. Herring, "Why and How to Design a Contingent Convertible Debt Requirement," 2011, 1-58.
http://fic.wharton.upenn.edu/fic/papers/11/11-41.pdf

[6.] CEIOPS, "Insurance Linked Securities Report," 2009, 4-23.
https://eiopa.europa.eu/fileadmin/tx_dam/files/publications/reports/CEIOPS-ILS-Report-Spring-2009.pdf

[7.] CEIOPS, "QIS5 Technical Specifications," 2010, 295-304.
http://ec.europa.eu/internal_market/insurance/docs/solvency/qis5/201007/technical_specifications_en.pdf

[8.] CEIOPS, "Task Force Report on the Liquidity Premium," 2010, 1-33.
https://eiopa.europa.eu/fileadmin/tx_dam/files/publications/submissionstotheec/20100303-CEIOPS-Task-Force-Report-on-the-liquidity-premium.pdf

[9.] CFO Forum, "Market Consistent Embedded Value Principles," 2009, 5-6.
http://www.cfoforum.nl/downloads/MCEV_Principles_and_Guidance_October_2009.pdf

[10.] Collender, Robert N., Forrest W. Pafenberg, and Robert S. Seiler, Jr., "Automatic Recapitalization Alternatives," FHFA Working Papers, 2010, 8-9.
http://www.fhfa.gov/webfiles/15543/workingpaper102r.pdf

[11.] COMMISSION SERVICES STAFF WORKING DOCUMENT, "POSSIBLE FURTHER CHANGES TO THE CAPITAL REQUIREMENTS DIRECTIVE," 2010, 18-19&23.
http://ec.europa.eu/internal_market/consultations/docs/2010/crd4/consultation_paper_en.pdf

[12.] Cooley, Thomas F., Viral V. Acharya, Matthew P. Richardson, Ingo Walter, "Regulating Wall Street: The Dodd-Frank Act and the New Architecture of Global Finance," 2010, 162-175.
Cox, Samuel H. and Hal W. Pedersen, "Catastrophe Risk Bonds," *North American Actuary Journal*, 2000, 56-82.

[13.] http://www.soa.org/library//journals/north-american-actuarial-journal/2000/october/naaj0010_5.pdf

[14.] Cox, S. H., J. R. Fairchild, and H. W. Pedersen, "Valuation of structured risk management products," *Insurance: Mathematics and Economics*, 2004, V34, 259-272.
Duffie, Darrell, and Kenneth J. Singleton, "Modeling Term Structures of Defaultable Bonds," *The Review of Financial Studies*, V12, Issue 4, 1999, 687-720.

[15.] European Commission, "International Accounting Standard 32 Financial Instruments: Presentation," 2011, 10.
http://ec.europa.eu/internal_market/accounting/docs/consolidated/ias32_en.pdf

[16.] THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION, "DIRECTIVE 2009/111/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL," Official Journal of the European Union, 2009, 97-98.
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:302:0097:0119:EN:PDF

[17.] "DIRECTIVE 2009/138/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II)," *Official Journal of the European Union*, 2009, 48-51.
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:335:0001:0155:en:PDF

[18.] FASB, "Statement of Financial Accounting Standards No. 133 Accounting for Derivative Instruments and Hedging Activities," 1998, 92.

[19.] Flannery, Mark J., "Stabilizing Large Financial Institutions with Contingent Capital Certificates," 2009, 5-21.
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1485689

[20.] FSB, "Reducing the moral hazard posed by systemically important financial institutions, FSB Recommendations and Time Lines," 2010, 1-3.
http://www.financialstabilityboard.org/publications/r_101111a.pdf

[21.] Garcia, Joao and Nicola Pede, "Pricing and Calibration of Contingent Capital with a Structural Approach,"

FitchSolutions Quantitative Analysis Special Report, 2011, 1-21.
[22.] Goldman Sachs, "Contingent capital Possibilities, problems and opportunities," 2011, 1-21. http://www.goldmansachs.com/our-thinking/topics/public-policy/regulatory-reform/contingent-capital.pdf
[23.] Hilscher, Jens and Alon Raviv, "Bank Stability and Market Discipline: The Effect of Contingent Capital on Risk Taking and Default Probability," 2011, 1-31. http://www.greta.it/credit/credit2011/PAPERS/Posters/29/7_Hilscher.pdf
[24.] Hull, J. and A. White: "One-Factor Interest Rate Models and the Valuation of Interest Rate Derivative Securities", Journal of Financial and Quantitative Analysis, 28 (June 1993a), 235-254.
[25.] Hull, J. and A. White: "Using Hull-White Interest Rate Trees", Journal of Derivatives, (Spring 1996), 26-36.
[26.] Jaimungal, Sebastian and Tao Wang, "Catastrophe options with stochastic interest rates and compound Poisson losses," *Insurance: Mathematics and Economics*, 2006, V38, 469-483.
[27.] Kashyap, Anil K., Raghuram G. Rajan and Jeremy C. Stein, "Rethinking Capital Regulation," 2008, 452-462. http://kansascityfed.org/publicat/sympos/2008/KashyapRajanStein.03.12.09.pdf
[28.] Koziol, Christian, and Jochen Lawrenz, "Contingent convertibles. Solving or seeding the next banking crisis?" 2011, 14-34.
[29.] Lanzenauer, Christoph Haehling Von, and William N. Lundberg, "The n-Fold Convolution of a Mixed Density and Mass Function," *ASTIN Bulletin International Actuarial Association*, V8, No. 1, 1974, 91-103. http://www.actuaries.org/LIBRARY/ASTIN/vol8no1/91.pdf
[30.] Lin, Shih-Kuwi and Chia-Chien Chang, "Catastrophe Equity Put in Markov Jump Diffusion Model," *ASTIN Colloquium*, 2007, 1-21. http://www.actuaries.org/ASTIN/Colloquia/Orlando/Papers/Lin1.pdf
[31.] Lin, Shih-Kuwi and Chia-Chien Chang, "Catastrophe Insurance Products in Markov Jump Diffusion Model," *ASTIN Colloquium*, 2007, 1-29. http://www.actuaries.org/ASTIN/Colloquia/Orlando/Papers/Lin2.pdf
[32.] Llyods Banking Group, "Exchange Offer and Publication of Exchange Offer Memorandum," 2009, 2-5. http://www.lloydsbankinggroup.com/media/pdfs/investors/2009/2009Nov3_LBG_Exchange_Offer_Publication_of_Exchange_Offer_Memo.pdf
[33.] Ma, N.-Y. and Feng Liu, "A novel analytical scheme to compute the n-fold convolution of exponential-sum distribution functions," *Applied Mathematics and Computation*, V158, 2004, 225–235. http://www.eng.utah.edu/~fliu/pdfs/Applied_math_comput_158_255_2004.pdf
[34.] Maes, Stan and Wim Schoutens, "Contingent capital: an in-depth discussion," 2010, 7. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1653877
[35.] Merton, Robert. C., "On the pricing of corporate debt: the risk structure of interest rates," *Journal of Finance*, V29, 1974, 449–470.
[36.] McDonald, Robert L., "Contingent Capital with a Dual Price Trigger," 2011, 1-32. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1553430
[37.] Millennium Consulting Services, LLC, "NAIC Regulatory Update," Volume V, Issue II, 2010, 10. http://www.mcs-nc.com/assets/pdfs/NAICnews_Aug10.pdf
[38.] Moody's., "Rating Considerations for Contingent Capital Securities," 2010, 1-6.
[39.] More, Jorge, Burt Garbow, and Ken Hillstrom "Minpack," 1980, 2. http://www.netlib.org/minpack/
[40.] OSFI, « Non-Viability Contingent Capital, » 2011, 1-7. http ://www.osfi-bsif.gc.ca/app/DocRepository/1/eng/guidelines/capital/advisories/nvcc_e.pdf
[41.] Pierre Hillion and Theo Vermaelen, "Death Spiral Convertibles," 2001, 1-8. http://jfe.rochester.edu/03026.pdf
[42.] Prescott, Edward Simpson, "Contingent Capital: The Trigger Problem," 2011, 5-15. http://richmondfed.org/publications/research/working_papers/2011/pdf/wp11-07.pdf
[43.] PricewaterhouseCoopers LLP, "Insurance Industry NAIC Meeting Notes," 2010, 12-13. http://www.pwc.com/en_US/us/insurance/assets/summernaic2010.pdf
[44.] Rubinstein, Mark and Eric Reiner, "Unscrambling the Binary Code", Risk 4, No. 8, (September 1991), 75-83
[45.] Sec. 165. "Enhanced supervision and prudential standards for nonbank financial companies supervised by the Board of Governors and certain bank holding companies," One Hundred Eleventh Congress of the United States of America, H.R.4174, 2010. http://www.sec.gov/about/laws/wallstreetreform-cpa.pdf

[46.] Spiegeleer, Jan De and Wim Schoutens, "Pricing Contingent Convertibles: A Derivatives Approach," 2011, 1-35.
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1795092
[47.] Squam Lake Working Group on Financial Regulation, "An Expedited Resolution Mechanism for Distressed Financial Firms: Regulatory Hybrid Securities," 2009, 2-5.
http://www.cfr.org/content/publications/attachments/Squam_Lake_Working_Paper3.pdf
[48.] "Standard & Poor's Ratings Services Criteria Regarding Contingent Capital Securities," 2009, 2.
[49.] Sundaresan, Suresh and Zhenyu Wang, "On the Design of Contingent Capital with Market Trigger," *Federal Reserve Bank of New York Staff Reports*, 2011, 11-14.
http://www.newyorkfed.org/research/staff_reports/sr448.pdf
[50.] Takahashi, Akihiko, Takao Kobayashi and Naruhisa Nakagawa, "Pricing Convertible Bonds with Default Risk: A Duffie-Singleton Approach," *The Journal of Fixed Income*, V11, No. 3, December 2001, 20-29.

Abbreviations and notations

Collect here in alphabetical order all abbreviations and notations used in the paper

AOF, ancillary own fund
BCBS, Basel Committee on Banking Supervision
CAB, capital access bond
CAR, capital adequacy ratio
CEP, catastrophe equity put
CCC, contingent capital certificates
CoCo bond, contingent convertible bond
CRD, capital requirements directive
CRNHR, cost of residual non hedgeable risks
CTE, conditional tail expectation
DTI, deposit-taking institution
FASB, Financial Accounting Standards Board
FSB, Financial Stability Board
GSIB, global systematically important bank
IAS, International Accounting Standards
ILWs, industry loss warranties
IFRS, International Financial Reporting Standards
MCEV, market consistent embedded value
MCR, minimum capital requirement
NVCC, non-variability contingent capital
OSFI, Office of the Superintendent of Financial Institutions Canada
RoE, return on equity
SCR, solvency capital requirement
SIFI, systematically important financial institution
VaR, value at risk
VBA, Visual Basic Application

Biography of the Author

Kailan Shang is a pricing actuary at Manulife Financial in Canada. Before that, he worked in the area of financial risk management and risk analytics in AIA. Years of actuarial and risk management experience has allowed him to get a broad exposure in the fields of economic capital, market-consistent embedded value, financial engineering, dynamic management options and policyholder behavior modeling, product development and management, financial reporting, dynamic solvency testing, and the like.

As an FSA, CFA, PRM, and SCJP, he is also an enthusiast of actuarial research through both volunteer works and funded research program. He participated in the LSMWP and was awarded the Emerging Issues Prize from the CAS for the paper titled "Loss Simulation Model Testing and Enhancement." He co-authored the paper "Risk Appetite: Linkage with Strategic Planning" sponsored by the Joint Risk Management Section of the CAS, CIA, and the SOA.

He can be reached at kailan_shang@manulife.com or klshang81@gmail.com.

An Actuarial Model of Excess of Policy Limits Losses

Neil M. Bodoff, FCAS, MAAA

Abstract

Motivation. Excess of policy limits (XPL) losses is a phenomenon that presents challenges for the practicing actuary.

Method. This paper proposes using a classic actuarial framework of frequency and severity, modified to address the unique challenge of XPL.

Results. The result is an integrated model of XPL losses together with non-XPL losses.

Conclusions. A modification of the classic actuarial framework can provide a suitable basis for the modeling of XPL losses and for the pricing of the XPL loss component of reinsurance contracts.

Keywords. Excess of Policy Limits. XPL. ERM. Modeling.

1. INTRODUCTION

Excess of policy limits (XPL) losses is a phenomenon that presents challenges for the practicing actuary. For example, exposure rating, one of the standard actuarial methods for pricing reinsurance layers, seems to be completely unworkable for the challenge of pricing XPL losses; yet often, an exposure rating approach to reinsurance pricing is the only method that the practicing actuary has at his disposal.

In this paper, I propose an approach that incorporates XPL into the classic actuarial framework of frequency, severity, and limited expected value (LEV) of claims. In this way, XPL will simply be part of a broader landscape of claims behavior, and can draw upon and seamlessly integrate with standard actuarial tools for incorporating the price of XPL losses into the pricing of reinsurance contracts. In addition, using the classic actuarial framework allows one to incorporate XPL losses into stochastic economic capital models that are used for insurer enterprise risk management (ERM) purposes.

1.1 Research Context

The actuarial literature has very limited discussion of actuarial approaches to modeling of excess of policy limits losses. I have found only one paper by Braithwaite and Ware [1], which remains a crucially important paper.

1.2 Objective

In this paper, I propose a framework that builds upon the work of Braithwaite and Ware yet differs in some ways.

There are two main reasons for this difference in approach. The first reason relates to aligning resources with need. XPL is an important actuarial problem but by no means the paramount problem typically facing actuaries. As a result, I would like to propose a reasonable methodology that is more practicable than the one proposed in Braithwaite and Ware. Whereas Braithwaite and Ware's model required the actuary to build an additional, freestanding size-of-loss curve to describe XPL, this paper proposes a methodology that simply extends one's existing size-of-loss curve, greatly simplifying the implementation.

The second reason that the proposed approach differs from Braithwaite and Ware is the need to quantify XPL losses in the context of a broader insurance portfolio; one ought to model and price for XPL in conjunction with other non-XPL losses. Braithwaite and Ware, discussing clash reinsurance treaties, focuses entirely on XPL losses. Yet the practitioner actuary often desires to price for XPL losses in working layer reinsurance; only a small percentage of losses will be XPL whereas the majority of losses will be non-XPL. The task, then, is to price these reinsurance layers for the XPL losses in a framework that aligns with traditional actuarial pricing methods. Similarly, another situation that requires modeling of XPL losses together with non-XPL losses is enterprise risk management (ERM), in which one seeks to model all the insurance risk of the company. Modeling requires an integrated framework that covers XPL and non-XPL losses together, which will be facilitated by the proposed new approach.

2. ACTUARIAL MODEL OF SIZE OF LOSS DISTRIBUTION WITH EXTENSION TO XPL

We begin with the classic actuarial framework for evaluating loss costs in layers with a focus on limited expected value (LEV). Following Clark [2], we can write that

X = random variable for size of loss

$F_X(x)$ = probability that random variable X, the size of loss, is less than or equal to x

$f_X(x)$ = probability density function, first derivative of F(x)

E[X] = expected value or average unlimited loss

E[X;k] = expected value of loss capped at k

The expected value of loss capped at an amount k can be defined as follows:

$$LEV(X,k) = E[X;k] = \int_0^k xf(x)dx + \int_k^\infty kf(x)dx \qquad (2.1)$$

$$LEV(X,k) = E[X;k] = \int_0^k xf(x)dx + k[1 - F(k)] \qquad (2.2)$$

2.1 Limited Expected Value (LEV)

Historically, actuaries needed to quantify the value of the average loss limited by the insurance policy; they adopted limited expected value (LEV) as the framework to calculate this value, under the assumption that a policy limit caps the insurance loss.

2.2 Incorporating XPL Losses

In light of our knowledge of XPL losses, we should revisit whether LEV is the ideal way to measure losses to an insurance policy. Let's describe the average loss accruing to an insurance policy as the Policy Limited Expected Value (PLEV). Until now, the implicit assumption has been that PLEV = LEV.

The phenomenon of XPL losses shows us, however, that the policy limit written in the insurance policy contract is not always potent in capping losses. Thus the identity function, PLEV = LEV, is not fully accurate.

What could be a paradigm for how to think about the phenomenon of XPL losses? I propose that we begin to think of the effectiveness of the policy limit as being subject to a random variable.

Let's define a random variable Z, which follows a Bernoulli distribution. This random variable can have a value of 1, or "success", with probability p, and can have a value of 0, "failure", with probability 1-p. When Z=1 we have "success" and the policy limit caps the insurance loss; when Z=0 we have "failure" and the policy limit does not cap the insurance loss and we have an XPL situation.

Now we can say that the Policy Limited Expected Value is:

$$PLEV(X,k,Z)=\int_0^k xf(x)dx+P(Z=1)*\int_k^\infty kf(x\mid Z=1)dx+P(Z=0)*\int_k^\infty xf(x\mid Z=0)dx \qquad (2.3)$$

Recalling that the probability that Z=1 is p and that Z=0 is 1-p, we write:

$$PLEV(X,k,Z)=\int_0^k xf(x)dx+p\int_k^\infty kf(x\mid Z=1)dx+(1-p)\int_k^\infty xf(x\mid Z=0)dx \qquad (2.4)$$

If we let x = k + (x-k) in the final integral, we can rewrite equation (2.4) is as follows:

$$PLEV(X,k,Z)=\int_0^k xf(x)dx+k[1-F(k)]+(1-p)\int_k^\infty (x-k)f(x\mid Z=0)dx \qquad (2.5)$$

One can say that on a fundamental level, equation (2.5) captures the approach crystallized in Braithwaite and Ware. The additional loss above and beyond the policy limit follows a different conditional probability density function than the initial size of loss distribution; as a result, the XPL loss component is a completely new entity that is grafted onto the non-XPL loss component.

3. A MORE PRACTICAL MODEL

How can we make this model more practical and easier to use? Let's revisit equation (2.4) and make some simplifying assumptions.

Let's assume that the probability density function above the policy limit is not conditional on whether or not an XPL scenario has been triggered. As explained in Braithwaite and Ware, the XPL situation arises when the policyholder is found liable for actual damage to a third party; the only question is whether or not the insurance company's conduct provides a basis for the courts to override the capping effect of the policy limit. Thus, this simplifying assumption should be reasonable for XPL (although perhaps not for extra-contractual obligations, ECO).

We can then substitute the unconditional f(x) into equation (2.4) by replacing the conditional f(x|Z=0) and f(x|Z=1) and rewrite equation (2.4) as follows:

$$PLEV(X,k,Z) = \int_0^k xf(x)dx + p\int_k^\infty kf(x)dx + (1-p)\int_k^\infty xf(x)dx \tag{3.1}$$

Thus we simply say that if random variable Z=1 we have a success and the policy limit caps the loss and if Z=0 we have a failure and the policy limit does not cap the loss. Unlike equation (2.5) and unlike the approach of Braithwaite and Ware, the XPL loss is not a completely new entity; rather, the XPL loss is simply an extension of the standard size-of-loss distribution that occurs when the policy limit's capping effect is ineffective. Such a framework would be much easier to work with when attempting to incorporate XPL losses.

3.1 Practical Applications: Insurance Risk Modeling

How can we apply the proposed paradigm of equation (3.1) in a practical way to achieve a tangible result? One possibility would be in a simulation environment.

3.1.1 Simulation Application #1: Collective Risk Model for Insurance Losses

Step #1: Define the size of loss distribution for an insurance policy or portfolio of policies on a gross of policy limit basis.

Step #2: Simulate individual losses and simulate the limit of the policy associated with each loss.

Step #3: For each loss, if the loss is greater than the policy limit, then simulate Z, a Bernoulli random variable. If Z=1, then cap the simulated loss at the policy limit. If Z=0, then do not cap the loss.

Notice that there is only one small new step here: rather than always capping the loss at the policy limit, let the capping be subject to the outcome of a random variable that reflects whether the policy limit will be effective at capping the loss or not.

3.1.2 Simulation Application #2: Catastrophe ("Cat") Modeling

The software vendors for cat modeling typically employ several steps in their calculations of the losses to an insurance portfolio for a given simulated cat event. After the software simulates a catastrophic ("cat") event, the software evaluates how the physical phenomenon affects the physical structures in its path. Then, in one of the final steps, the software overlays the insurance policy's contractual terms to achieve the financial loss to the company. Within this simulation environment, the final step could evolve away from the current deterministic view of the policy limit and towards a stochastic view of the policy limit. Moreover, one could consider correlating the individual probabilities that the policy limits fail; the correlation could depend upon geographical location and legal jurisdiction, among other factors. An approach to cat modeling simulations that treats policy limit capping of losses as a probable but not definite outcome would be more realistic and would show more severe risk metric output than current models.

3.2 Reinsurance Pricing

A second practical application of the proposed paradigm of equation (3.1) could be reinsurance pricing.

Recall that traditional exposure rating is viewed as not producing loss cost indications that encompass XPL. After all, XPL losses by definition exceed the policy limit and thus exceed the exposure; how could exposure rating possibly incorporate XPL within its framework?

Let's revisit equation (3.1):

$$PLEV(X,k,Z) = \int_0^k xf(x)dx + p\int_k^\infty kf(x)dx + (1-p)\int_k^\infty xf(x)dx \tag{3.1}$$

If we multiply the first term on the right side of equation (3.1) by 1 and let 1 = p + 1 – p and rearrange terms, we can rewrite equation (3.1) as follows:

$$PLEV(X,k,Z) = p[\int_0^k xf(x)dx + \int_k^\infty kf(x)dx] + (1-p)[\int_0^k xf(x)dx + \int_k^\infty xf(x)dx] \tag{3.2}$$

This is also the same as the following:

$$PLEV(X,k,Z) = p[\int_0^k xf(x)dx + \int_k^\infty kf(x)dx] + (1-p)[\int_0^\infty xf(x)dx] \tag{3.3}$$

And:

$$PLEV(X,k,Z) = p(LEV(X,k)) + (1-p)E[X] \tag{3.4}$$

Equations (3.3) and (3.4) demonstrate that in the presence of XPL losses, we have a loss severity that has probability p of being limited by the policy limit and probability (1-p) of not being limited by the policy limit.

We can use this framework to calculate expected layer loss for excess-of-loss reinsurance exposure rating.

Following Clark, for each policy we want to calculate the exposure factor, i.e. the percentage of the policy's total loss that is covered by the reinsurance layer.

$$Exposure\ Factor = \frac{layer\ loss}{total\ loss} \tag{3.5}$$

Now let's calculate the layer loss.

Layer loss = Loss limited at the top of the reinsurance layer – loss limited at the bottom of the reinsurance layer (3.6)

Here, we have a probability p that the policy limit will cap the loss and a 1-p probability that the policy limit will not cap the loss. While these probabilities apply to the primary policy, we assume that they do not apply at all to the reinsurance limit and attachment point.

Thus, when estimating the loss limited by the top of the reinsurance layer, we have a probability p that the loss will be capped by the lesser of the policy limit and the top of the reinsurance layer; we also have a probability 1-p that the loss will be capped solely by the top of reinsurance layer, with no application of the policy limit.

Loss limited at top of reinsurance layer = p * LEV (X, min(policy limit, reinsurance exit point)) + (1-p) * LEV (X, reinsurance exit point)	(3.7)

Note: Reinsurance exit point = reinsurance attachment point + reinsurance limit

Similarly, when estimating the loss limited by the bottom of the reinsurance layer, we have a probability p that the loss will be capped by the lesser of the policy limit and the bottom of the reinsurance layer; we also have a probability 1-p that the loss will be capped solely by the bottom of reinsurance layer.

Loss limited at bottom of reinsurance layer = p * LEV (X, min(policy limit, reinsurance attachment point)) + (1-p) * LEV (X, reinsurance attachment point)	(3.8)

Thus:

Layer loss = p * LEV (X, min(policy limit, reinsurance exit point)) + (1-p) * LEV (X, reinsurance exit point) – {p * LEV (X, min(policy limit, reinsurance attachment point)) + (1-p) * LEV (X, reinsurance attachment point)}	(3.9)

Thus:

Layer loss = p * traditional exposure rating layer LEV subject to primary policy limit + (1-p) * layer LEV not subject to primary policy limit	(3.10)

Having calculated the layer loss, which is the numerator of the exposure factor, we now need to calculate the denominator, the policy's total loss.

Recall that the exposure factor produces layer loss by multiplying the policy's total loss; total loss is usually calibrated based on policy premium multiplied by an Expected Loss Ratio (ELR). Therefore, whether or not the ELR was calculated to include a provision for XPL losses will affect how one ought to calculate the denominator of the exposure factor.

For our discussion, let's proceed under the assumption that the ELR does not include a provision for XPL loss. As a result, when calculating the "total loss" for the denominator of the exposure factor, we will calculate it based only on non-XPL losses.

Denominator of Exposure Factor = Same as traditional exposure rating = Policy total loss excluding XPL = LEV(X, policy limit)	(3.11)

Then, combining equations (3.9) and (3.11), we derive:

Exposure Factor = [p * LEV (X, min(policy limit, reinsurance exit point)) + (1-p) * LEV (X, reinsurance exit point) – {p * LEV (X, min(policy limit, reinsurance attachment point)) + (1-p) * LEV (X, reinsurance attachment point)}] / LEV(X, policy limit)	(3.12)

Or, more simply, combining equations (3.10) and (3.11), we derive:

Exposure Factor = [p * traditional exposure rating layer LEV subject to primary policy limit + (1-p) * layer LEV not subject to primary policy limit] / traditional exposure rating ground up LEV capped at policy limit	(3.13)

3.2.1 Reinsurance Pricing: Numerical Example

Now let's do a numerical example of the proposed algorithm. The goal is to generate layer loss costs via exposure rating that include a loss provision for XPL losses.

First, let's stipulate some hypothetical numerical values for our policy limits distribution:

Exhibit 1

1	2	3
Policy Limit	% of premium	ELR%
50,000	1.0%	65.0%
100,000	1.0%	65.0%
500,000	2.0%	65.0%
1,000,000	80.0%	65.0%
2,000,000	10.0%	65.0%
3,000,000	1.0%	65.0%
4,000,000	1.0%	65.0%
5,000,000	3.0%	65.0%
10,000,000	1.0%	65.0%

We also need values for our size-of-loss severity curve:

Exhibit 2

Item #	Description	Value
1	Curve	Pareto
2	Theta	50,000
3	Alpha	1.50

Finally, we need to input parameter values for probability p that a policy limit will successfully cap losses and 1-p that the policy limit will not cap losses; the values may vary for each policy. Here we select a simple parameter structure in which all the policies in our limits table have the same value for p.

Exhibit 3

	p	1-p
All Policy Limits < $25M	99%	1.00%
Policy Limit = $25M	100%	0.00%

We now apply the proposed methodology to the numerical values to produce the following output in Exhibit 4.

Exhibit 4

1	2	3	4	5	6
			Layer Losses as % of total ground up losses	Layer Losses as % of total ground up losses	Implied Loading for XPL
Layer	Limit	Attachment	Traditional Exposure Rating	Proposed Method Including XPL	Proposed / Traditional - 1
1	500,000	-	88.420%	88.440%	0.023%
2	500,000	500,000	10.067%	10.074%	0.072%
3	1,000,000	1,000,000	1.150%	1.219%	5.989%
4	3,000,000	2,000,000	0.333%	0.403%	21.057%
5	5,000,000	5,000,000	0.031%	0.068%	119.369%
6	15,000,000	10,000,000	0.000%	0.033%	#N/A
Total			100.000%	100.237%	0.237%

Column 6 of Exhibit 4 shows the "loading factor" for each layer loss attributable to XPL. What is notable about this output is that choosing one simple value for p creates layer loading factors for XPL that are different for the various layers. Also, these loading factors for XPL would be different for other portfolios with different policy limits distributions, even with no change in the underlying value of the p parameters.[1]

[1] A copy of the Microsoft Excel workbook with the supporting calculations is available from the author upon request.

4. CONCLUSIONS

In this paper, I propose an actuarial paradigm for describing excess of policy limits (XPL) losses. The central idea is that one can envision a random variable governing the application of the policy limit; most of the time the policy limit is enforced as it is written in the insurance contract, whereas other times the policy limit is superseded. This paradigm is quite parsimonious; therein lies its attractiveness. At the same time, this simple framework can generate nuanced, differentiated, useful, and non-obvious output information for practicing actuaries. One practical application would be to incorporate XPL losses into actuarial exposure rating estimates for casualty excess-of-loss reinsurance layers; the output values vary based on the attachment point and limit of the reinsurance layer being priced as well as the granular policy limits usage of the particular insurance portfolio under review. A second practical application would be to incorporate XPL losses in a simulation environment such as commercial software for estimating losses arising from natural catastrophes; envisioning policy limits as being random variables can affect the cat modeling and thus the critical risk metrics of an insurer's portfolio.

5. REFERENCES

[1] Braithwaite, Paul, and Bryan C. Ware, "Pricing Extra-Contractual Obligations and Excess of Policy Limits Exposures in Clash Reinsurance Treaties", *CAS Forum*, Spring 1997.
[2] Clark, David R., "Basics of Reinsurance Pricing", *CAS Study Note*, 1996.

Biography of the Author

Neil Bodoff is Executive Vice President at Willis Re Inc. He is a Fellow of the Casualty Actuarial Society.

Contact the author at neil.bodoff@willis.com and neil_bodoff@yahoo.com

Classifying the Tails of Loss Distributions

Leigh J. Halliwell, FCAS, MAAA

Abstract. Of the several classifications which actuaries have proposed for the heaviness of loss-distribution tails, none has been generally accepted. Here we will show that the ultimate settlement rate, or asymptotic failure rate, provides a natural tripartite division into light, medium, and heavy tails. We prove that all the positive moments of light- and medium-tailed distributions are finite. Within the heavy-tailed distributions, we will define very heavy-tailed and super heavy-tailed, and we will explain how the power and exponential transformations are the basis for these subdivisions. An appendix relates extreme value theory to our findings.

Keywords: loss distribution, ultimate settlement rate, power transform, exponential transform, extreme value theory

1. INTRODUCTION

Many actuaries are as fascinated with the "heaviness" of loss-distribution tails as chemists are with heavy elements and as physicists are with heavy particles. However, unlike those scientists, with their periodic table and standard particle theory, "actuarial scientists" have no generally accepted standard of tail comparison. In this paper we will propose one that gives every indication of being natural, comprehensive, and insightful. To outline our progress, after briefly defining in Section 2 what constitutes a loss distribution, in Section 3 we will introduce the ultimate settlement rate and derive the settlement rates of several familiar distributions. Then in Section 4 we will show how the most basic transformation, a change of a distribution's scale parameter, provides the basis for a division of loss distributions into light-, medium-, and heavy-tailed. An immediate benefit from this is a proof in Section 5 that all the positive moments of light- and medium-tailed distributions are finite. Infinite moments are a sufficient, but not a necessary condition, for being heavy-tailed. Section 6 takes up the next logical transformation, the power transformation, and will show its effect on the tail class of a distribution. A symmetric "multiplication" table there, showing the medium-tailed distribution to be like an identity element among distributions, will be crucial to the following sections. Section 7 contains an abstract examination into the results so far, finishing with a diagram that will make memorable the classification schema. In Section 8 we will treat the next logical transformation, the exponential, which is the key to loss-tail heaviness. Then in Section 9 we will treat the moments of exponentially transformed random variables, vindicating the power of this classification by the results. Section 10 is a brief treatment of two other transformations, inverting and mixing. Finally, before concluding, in Section 11 we will show that the classification is indefinitely expandable, encompassing ever more distant realms of heavy and light tails. An appendix will fit extreme value theory into the classification schema.

2. LOSS DISTRIBUTIONS DEFINED

For the purposes of this paper *X* is a loss distribution[1] if its survival function $S_X(x) = Prob[X > x]$ has the following properties:

(i) $S_X(0) = 1$

(ii) $S_X(x) > 0$

(iii) For all $x_1 < x_2$, $S_X(x_1) \geq S_X(x_2)$. And there exists some ξ such that for all $\xi < x_1 < x_2$, $S_X(x_1) > S_X(x_2)$

(iv) $\lim_{x \to \infty} S_X(x) = 0$

(v) For all *x* greater than some ξ, $S''_X(x) > 0$

Although these properties are standard, some commentary will be helpful. Property (i) implies that *X* must be positive; in particular, there is no probability mass at zero. So this definition disqualifies the Tweedie distribution (Meyers [5]; cf. Footnote 15). The property provides for $1/X$ to be a loss distribution, which we deem desirable.[2] Property (ii) requires the tail of a loss distribution to be infinite. We are not interested in classifying tails of finite distributions; they might as well be "no-tailed." Property (iii) requires for the survival function never to increase, and beyond some point for it strictly to decrease. Property (iv) precludes any probability that *X* might be infinite. Though we will often encounter limits to infinity, infinity is not a real number. This property is allied with the first, for if somehow $Prob[X = \infty] > 0$, then the inverse $1/X$ would have a probability mass at zero. And property (v) demands beyond some point for the survival function to be concave upward. Of course, for the second derivative to exist the first derivative must also exist. By implication, the left and right derivatives must be equal. This property ensures that at some point the survival function "settles down." Thereafter there will be no more discrete jumps or probability masses, no more vertices or corners, and no more undulations or inflections. Having defined a loss distribution, we name the set of all loss distributions Ξ.

3. THE ULTIMATE SETTLEMENT RATE

[1] More accurately, *X* is a "loss random variable," whose probability obeys a "loss distribution." But since "loss random variable" sounds odd, we will use 'random variable' and 'distribution' interchangeably.

[2] We also desire $E[X^0]$ to equal unity, which would fail if there were any probability of the indeterminate 0^0. Cf. Section 5, esp. Footnote 6.

Of the several ways described by Klugman [4, pp. 86-92] and Corro [1] by which to compare the tails of loss distributions, we believe the best to be the "asymptotic failure rate" [4, p. 87] or "ultimate settlement rate" [1, p. 451]. Since $S_X(x) = Prob[X > x]$ is the probability for *X* to "survive" at least until *x*, we may think of *X* as being subject to a force of mortality $\lambda_X(x) = -\frac{d\,ln\,S_X(x)}{dx} = -\frac{1}{S_X(x)}\frac{dS_X(x)}{dx} = \frac{f_X(x)}{S_X(x)}$. Corro's ultimate settlement rate is $\tau_X = \lim_{x \to \infty} \lambda_X(x)$. Just as an account compounding at a higher interest rate will eventually overtake an account compounding at a lower rate, regardless of their current positive balances, so too if $\tau_X < \tau_Y$, $\frac{S_X(x)}{S_Y(x)}$ will grow infinitely large with *x*, i.e., $\lim_{x \to \infty} \frac{S_X(x)}{S_Y(x)} = \infty$ or $\lim_{x \to \infty} \frac{S_Y(x)}{S_X(x)} = 0$.[3] But, following Klugman [4, p. 88], from L'Hôpital's rule we may express τ_X in terms of the probability density function:

$$\begin{aligned}
\tau_X &= \lim_{x \to \infty} \lambda_X(x) \\
&= \lim_{x \to \infty} \frac{f_X(x)}{S_X(x)} \\
&= \lim_{x \to \infty} \frac{f_X(x)}{\int_{u=x}^{\infty} f_X(u)} \qquad \text{a } \frac{0}{0} \text{ form} \\
&= \lim_{x \to \infty} \frac{f'_X(x)}{(-f_X(x))} \\
&= -\lim_{x \to \infty} \frac{d\,ln\,f_X(x)}{dx}
\end{aligned}$$

This does not mean that $\lambda_X(x) = -\frac{d\,ln\,f_X(x)}{dx}$; it is only true in the limit as $x \to \infty$. Of course, for $0 \le \xi \le x$, where ξ is the "settling down" point required by property (v):

$$S_X(x) = S_X(\xi)e^{-\int_{u=\xi}^{x} \lambda_X(u)du}$$

Let us look at the settlement rates of some well known distributions. If $X \sim Gamma(\alpha, \theta)$, or equivalently $X/\theta \sim Gamma(\alpha, 1)$, then $f_X(x) = \frac{1}{\Gamma(\alpha)} e^{-\frac{x}{\theta}} \left(\frac{x}{\theta}\right)^{\alpha-1} \frac{1}{\theta}$, for positive α. If $-\alpha < k$, $E[X^k] = \theta^k \frac{\Gamma(\alpha + k)}{\Gamma(\alpha)}$. Now:

---

[3] This implies an infinite right tail for *X* (property ii). Although in order for $S_X(x)$ to reach zero the force of mortality $\lambda_X(x)$ must become infinite, once $S_X(x)$ "flatlines" at zero, $\lambda_X(x) = 0/0$. It is meaningless to speak of the growth (or mortality) rate of something whose quantity is zero; a zero balance in a bank account remains zero at any interest rate. Property (v) guarantees the existence of $\lambda_X(x)$ far enough out, as well as for $\lim_{x \to \infty} \lambda_X(x)$ either to converge to a non-negative real number or to diverge to positive infinity.

$$\tau_{Gamma(\alpha,\theta)} = -\lim_{x\to\infty}\frac{d\,ln\,f_X(x)}{dx} = -\lim_{x\to\infty}\frac{d\left(-\frac{x}{\theta}+(\alpha-1)ln\,x\right)}{dx} = -\lim_{x\to\infty}\left(-\frac{1}{\theta}+\frac{\alpha-1}{x}\right) = \frac{1}{\theta}$$

The ultimate settlement rate of a Gamma-distributed random variable depends only on its scale parameter θ. But the force of mortality of the exponential random variable is $\lambda_{Gamma(1,\theta)}(x) = \frac{1}{\theta}$. For this reason it is legitimate to say that *far enough out in the tail, every gamma distribution looks like an exponential distribution*. Compare this with the right tail of a normal distribution:[4]

$$\tau_{N(\mu,\sigma^2)} = -\lim_{x\to\infty}\frac{d\,ln\,f_X(x)}{dx} = -\lim_{x\to\infty}\frac{d\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right)}{dx} = \lim_{x\to\infty}\left(\frac{x-\mu}{\sigma^2}\right) = \infty$$

So the right tail of the normal distribution is "lighter" than that of the gamma.

For the inverse-gamma random variable, $(X/\theta)^{-1} \sim Gamma(\beta,1)$, or $X/\theta \sim 1/Gamma(\beta,1)$, where $\beta > 0$. Its density function is $f_X(x) = \frac{1}{\Gamma(\beta)}e^{-\frac{\theta}{x}}\left(\frac{\theta}{x}\right)^{\beta+1}\frac{1}{\theta}$, and $E[X^k] = \theta^k\frac{\Gamma(\beta-k)}{\Gamma(\beta)}$, for $k < \beta$. As for its ultimate settlement rate:

$$\tau_{InvGamma(\beta,\theta)} = -\lim_{x\to\infty}\frac{d\,ln\,f_X(x)}{dx} = -\lim_{x\to\infty}\frac{d\left(-\frac{\theta}{x}-(\beta+1)ln\,x\right)}{dx} = -\lim_{x\to\infty}\left(\frac{\theta}{x^2}-\frac{\beta+1}{x}\right) = 0$$

So the inverse-gamma is "heavier-tailed" than the gamma distribution, since $0 < 1/\theta$. One might have surmised this from the non-existence of its positive moments greater than or equal to β.

If X is a lognormal random variable, then $ln\,X \sim N(\mu,\sigma^2)$ and $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2}\frac{(ln\,x-\mu)^2}{\sigma^2}}\frac{1}{x}$. For all real *k*, $E[X^k] = e^{k\mu+k^2\sigma^2/2}$. All the moments of the lognormal random variable exist, even the negative ones. Its ultimate settlement rate is:

$$\tau_{LogNorm(\mu,\sigma^2)} = -\lim_{x\to\infty}\frac{d\,ln\,f_X(x)}{dx} = -\lim_{x\to\infty}\frac{d\left(-\frac{1}{2}\frac{(ln\,x-\mu)^2}{\sigma^2}-ln\,x\right)}{dx} = \lim_{x\to\infty}\left(\frac{ln\,x-\mu}{\sigma^2}\cdot\frac{1}{x}+\frac{1}{x}\right) = 0$$

Hence, the lognormal distribution is heavier-tailed than the gamma. Although its settlement rate equals the inverse-gamma's, the existence of all its moments implies that it is not as heavy-tailed as

[4] The normal distribution with its infinite left tail is not a loss distribution. But we may still calculate the ultimate settlement rate of its right tail. Alternatively, we could also consider the right tail of the absolute value of the standard normal distribution (i.e., $X/\theta \sim |N(0,1)|$) and arrive at the same result (cf. Footnote 10).

the inverse gamma.

Last, let X be a generalized-Pareto random variable. In our parameterization this will mean that $\frac{X}{\theta} \sim \frac{Gamma(\alpha,1)}{Gamma(\beta,1)}$, where the two gamma random variables are independent. The distribution function for X is $f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\left(\frac{x}{x+\theta}\right)^{\alpha-1}\left(\frac{\theta}{x+\theta}\right)^{\beta-1}\frac{\theta}{(x+\theta)^2}$, and its moments are $E\left[X^k\right] = \theta^k \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}\frac{\Gamma(\beta-k)}{\Gamma(\beta)}$, for $-\alpha < k < \beta = (-\alpha < k) \cap (k < \beta)$. Division by the $Gamma(\beta,1)$, or multiplication by the $InvGamma(\beta,1)$ random variable, places a positive limit on *k*. Its ultimate settlement rate is:

$$\tau_{GenPareto(\alpha,\beta,\theta)} = -\lim_{x\to\infty}\frac{d\,ln\,f_X(x)}{dx} = -\lim_{x\to\infty}\frac{d((\alpha-1)ln\,x - (\alpha+\beta)ln(x+\theta))}{dx} = \lim_{x\to\infty}\left(\frac{\alpha-1}{x} - \frac{\alpha+\beta}{x+\theta}\right) = 0$$

Again, this is the same rate as the inverse-gamma's and the lognormal's. But the domain of its positive moments makes its tail like the inverse-gamma's. The non-existence of negative moments is relevant only to the tail of the inverse distribution.

To conclude this section, the tail of the normal distribution is lighter than the tail of the gamma distribution, which is lighter than the tails of the lognormal, inverse-gamma, and generalized-Pareto distributions, even as $\infty > \theta > 0$. The non-existence, or infinitude, of positive moments hints at secondary orderings within the last three distributions.

4. THE ULTIMATE SETTLEMENT RATE UNDER A SCALE TRANSFORMATION

In the previous section we saw that the ultimate settlement rate of a gamma random variable is the inverse of its scale parameter. Here we will generalize, and form the basis for classifying the (right) tails of loss distributions.

If X is a random variable and θ a positive constant, the scale transformation of X is the random variable $Y/\theta = X$. Accordingly:

$$S_Y(u) = Prob[Y > u] = Prob[Y/\theta > u/\theta] = Prob[X > u/\theta] = S_X(u/\theta)$$

Hence:

$$\tau_Y = -\lim_{u\to\infty}\frac{d\,ln\,S_Y(u)}{du} = -\lim_{u\to\infty}\frac{d\,ln\,S_X(u/\theta)}{du} = -\lim_{u/\theta\to\infty}\frac{d\,ln\,S_X(u/\theta)}{d(u/\theta)}\frac{1}{\theta} = \frac{\tau_X}{\theta}$$

A scale transformation should not be the basis for tail class; in fact, most loss distributions are parameterized to include one "scale" parameter along with one or more "shape" parameters.[5] Because $0/\theta = 0$, $+/\theta = +$, and $\infty/\theta = \infty$, there are only three essential values for ultimate settlement rates, zero (0), positive (+), and infinity (∞), with the ordering, $0 < + < \infty$. Since smaller τ_X means heavier tail, we will classify a loss distribution as light-, medium, or heavy-tailed according as τ_X is ∞, +, or 0. The meaning of the symbols in the partition $\Xi = \Xi_0 \cup \Xi_+ \cup \Xi_\infty$ should be obvious. This is the gist of our classification; the rest of the paper merely draws out its implications.

[5] The scale parameter bears the unit of the random variable Y, so Y/θ is unitless, or a pure number. Most accurate is to divide each random variable by a parameter and to relate them by a unitless factor (called a "scalar"), as in $Y/\theta_1 = \eta(X/\theta_2)$. This is a safeguard in the physical sciences, but here it would be stilted. Cf. Footnote 13.

5. POSITIVE MOMENTS AND THE ULTIMATE SETTLEMENT RATE

In Section 3 we found loss distributions whose ultimate settlement rates spanned the range of zero, positive, and infinity. The normal distribution is light-tailed; the gamma distribution is medium-tailed; and the inverse-gamma, lognormal, and generalized-Pareto distributions are heavy-tailed. We also noted that some of the moments of the inverse-gamma and generalized-Pareto random variables were infinite. In this section we will prove that all the positive moments of light- and medium-tailed random variables are finite. But before that we will prove a partitioning lemma about non-negative moments, viz., that if $E[X^l]$ is finite for $0 < l$, then $E[X^k]$ is finite for $0 < k < l$.

If $X > 0$, then $X^0 = 1$.[6] Because according to property (i) $Prob[X > 0] = S_X(0) = 1$, $Prob[X^0 = 1] = 1$ and so $E[X^0] = E[1] = 1$. In words, the zeroth moment of a loss distribution exists and equals unity. Next consider $E[X^k]$ for $k > 0$. Because over this range of integration x^k is non-negative:

$$E[X^k] = \int_{x=0}^{\infty} x^k dF_X(x) = \int_{x=0}^{1} x^k dF_X(x) + \int_{x=1}^{\infty} x^k dF_X(x) \leq 1 + \int_{x=1}^{\infty} x^k dF_X(x)$$

So whether or not $E[X^k]$ is finite depends on $\int_{x=1}^{\infty} x^k dF_X(x)$. But for $x \geq 1$ and $k < l$, $1 \leq x^k \leq x^l$. So, if $\int_{x=1}^{\infty} x^l dF_X(x)$ converges, then so to does $\int_{x=1}^{\infty} x^k dF_X(x)$. Likewise, if $\int_{x=1}^{\infty} x^k dF_X(x)$ diverges, so too does $\int_{x=1}^{\infty} x^l dF_X(x)$. Therefore, for $0 < k < l$, if $E[X^l]$ is finite, so too is $E[X^k]$. And if $E[X^k]$ is infinite, so too is $E[X^l]$. The existence or non-existence of moments partitions the non-negative real numbers into two subsets. The lower partition is not empty, since it includes zero. The upper partition is empty when all the positive moments converge.

To return to the theorem of this section, let the distribution of *X* be light- or medium-tailed. So $\tau_X = \lim_{x\to\infty} \lambda_X(x) > 0$. And let $\rho = \tau_X/2$, if τ_X is finite; let $\rho = 1$, if it is infinite. In either case, $\rho > 0$ and there exists a $\xi > 0$ such that for all $x \geq \xi$, $\rho < \lambda_X(x)$. So for all $x \geq \xi$:

$$S_X(x) = S_X(\xi) e^{-\int_{u=\xi}^{x} \lambda_X(u)du} \leq S_X(\xi) e^{-\int_{u=\xi}^{x} \rho du} = S_X(\xi) e^{-\rho(x-\xi)}$$

[6] The form 0^0 is undefined, even as $0^0 = e^{ln 0 \cdot 0} = e^{-\infty \cdot 0}$. Corro's "convention" that $0^0 = 1$ [1, p. 453] is equivalent to the convention that $\infty \cdot 0 = 0/0 = 0$. This convention is wired into the arithmetic of some programming languages (e.g., APL and J. R is inconsistent: $0^0 = 1$, but $0/0$ is undefined). However, such conventions should not be placed on undefined, or indeterminate, forms, since in limiting cases they may assume different values.

By integration by parts one can show that $E[h(X)] = h(0) + \int_{x=0}^{\infty} S_X(x)dh(x)$.[7] So, for positive *k*:

$$E[X^k] = 0^k + \int_{x=0}^{\infty} S_X(x)dx^k = \int_{x=0}^{\infty} S_X(x)dx^k = \int_{x=0}^{\xi} S_X(x)dx^k + \int_{x=\xi}^{\infty} S_X(x)dx^k$$

$$\leq \int_{x=0}^{\xi} 1 \cdot dx^k + \int_{x=\xi}^{\infty} S_X(\xi)e^{-\rho(x-\xi)}dx^k$$

Finally, we simplify the inequality:

$$\begin{aligned} E[X^k] &\leq \int_{x=0}^{\xi} 1 \cdot dx^k + \int_{x=\xi}^{\infty} S_X(\xi)e^{-\rho(x-\xi)}dx^k \\ &= \xi^k + kS_X(\xi)e^{\rho\xi} \int_{x=\xi}^{\infty} e^{-\rho x}(\rho x)^{k-1} d(\rho x) \cdot \rho^{-k} \\ &\leq \xi^k + kS_X(\xi)e^{\rho\xi} \int_{\rho x=0}^{\infty} e^{-\rho x}(\rho x)^{k-1} d(\rho x) \cdot \rho^{-k} \\ &= \xi^k + kS_X(\xi)e^{\rho\xi}\,\Gamma(k)/\rho^k \end{aligned}$$

Thus we prove that $E[X^k]$ is not infinite. As a result, we know that all the positive moments of light- and medium-tailed distributions exist.

This converse ("Not all the positive moments of heavy-tailed distributions exist.") is not true, for the lognormal is heavy-tailed, yet all its moments exist. But a random variable that lacks even one positive moment is heavy-tailed. This suggests a subclass of the heavy-tailed distributions Ξ_0. Those lacking in positive moments are heavier-tailed than those not lacking. And the heaviest of many heavier-tailed distributions is the one with fewest positive moments (or the one with the most infinite moments).[8] But the next section will provide a better subclassification.

6. THE ULTIMATE SETTLEMENT RATE UNDER A POWER TRANSFORMATION

In Section 4 we found the tail classification of a random variable to be invariant to a scale transformation; more accurately, we devised that classification for it to be invariant. But just as

[7] For details cf. Halliwell [3, Appendix A]

[8] Are there distributions so heavy-tailed that they have no positive moments? In a Section 9 we will prove that there are such distributions. However, it seems that their worth is purely theoretical.

Klugman [4, pp. 92-93] advances from scale transformations to power transformations, so too will we in this section.

Our form of the power transformation is $\frac{Y}{\theta} = X^{\gamma}$, for positive γ.[9] The equation $E[Y^k] = \theta^k E\left[\left(\frac{Y}{\theta}\right)^k\right] = \theta^k E[(X^{\gamma})^k] = \theta^k E[X^{k\gamma}]$ puts the moments of *X* and *Y* into a one-to-one correspondence. Thus, distributions with infinite positive moments remain heavy-tailed under a power transformation. But how other distributions power-transform requires the following analysis.

Since x^{γ} strictly increases:

$$S_Y(u) = Prob[Y > u] = Prob[X^{\gamma} > u/\theta] = Prob\left[X > (u/\theta)^{1/\gamma}\right] = S_X\left((u/\theta)^{1/\gamma}\right)$$

Therefore:

$$\begin{aligned}
\tau_Y &= -\lim_{u\to\infty} \frac{d\ln S_Y(u)}{du} \\
&= -\lim_{u\to\infty}\left\{\frac{d\ln S_X\left((u/\theta)^{1/\gamma}\right)}{d(u/\theta)^{1/\gamma}}\frac{d(u/\theta)^{1/\gamma}}{du}\right\} \\
&= \lim_{\substack{u\to\infty \\ v(u)\to\infty}}\left\{-\frac{d\ln S_X(v)}{dv}\cdot\frac{1}{\gamma}(u/\theta)^{1/\gamma-1}\Big/\theta\right\} \\
&= \frac{1}{\gamma\theta}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot\left(v^{\gamma}\right)^{1/\gamma-1}\right\} \\
&= \frac{1}{\gamma\theta}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot v^{1-\gamma}\right\}
\end{aligned}$$

Now $\lim_{v\to\infty}\lambda_X(v) = \tau_X$. And $\lim_{v\to\infty} v^{1-\gamma} = \begin{cases} \infty & 0<\gamma<1 \\ 1 & \gamma = 1 \\ 0 & \gamma > 1 \end{cases}$. By the product rule we can express τ_Y in the following three-valued multiplication table (so $1/\gamma\theta$ may be ignored):

[9] Klugman [4, p. 93] uses 'τ' for the exponent ($Y = X^{1/\tau}$); we use 'γ' to avoid confusion with Corro's ultimate settlement rate τ. We also invert the exponent, because we believe it easier to see that $\gamma < 1$ thins the tail (taking a root) and $\gamma > 1$ thickens it (raising to a power).

$\tau_Y \quad \left(\frac{Y}{\theta} = X^{\gamma}\right)$	*Thinner* $0<\gamma<1$	$\gamma=1$	*Thicker* $\gamma>1$
Light: $\tau_X = \infty$	∞	∞	$\infty \cdot 0$
Medium: $\tau_X = +$	∞	$+$	0
Heavy: $\tau_X = 0$	$0 \cdot \infty$	0	0

Most obvious is the sensitivity of the medium-tailed random variable: the slightest exponent $\gamma = 1 \pm \varepsilon$ knocks it off the medium ridge into light or heavy valleys, from which the inverse exponent can restore it. For example, if *X* is medium-tailed, then $Y = X^2$ is heavy-tailed. And if $Y = X^{0.5}$, then *Y* is light-tailed.[10] So by power transformation, a medium-tailed distribution can become either heavy or light. But because $\left(Y^{\gamma}\right)^{1/\gamma} = Y = \left(Y^{1/\gamma}\right)^{\gamma}$, power transformations are invertible. By repeated transformations and inversions, one can cycle a medium-tailed distribution through all three types; e.g., $X \to \sqrt{X} \to \left(\sqrt{X}\right)^4 = X^2 \to \sqrt{X^2} = X$ is a three-stop roundtrip from medium to light to heavy and back to medium.

7. SET-THEORETIC PRESENTATION AND DIAGRAM OF RESULTS SO FAR

Define $PT[X;\gamma,\theta]$, the power transformation of random variable *X* with positive parameters γ and θ, as the distribution *Y* such that $\frac{Y}{\theta} = X^{\gamma}$. Therefore, $PT[X;\gamma,\theta] = \theta X^{\gamma}$. A compound power transformation reduces to a simple one:

$$\begin{aligned} PT\left[PT[X;\gamma_1,\theta_1];\gamma_2,\theta_2\right] &= \theta_2\left(PT[X;\gamma_1,\theta_1]\right)^{\gamma_2} \\ &= \theta_2\left(\theta_1 X^{\gamma_1}\right)^{\gamma_2} \\ &= \theta_2\theta_1^{\gamma_2} X^{\gamma_1\gamma_2} \\ &= PT\left[X;\gamma=\gamma_1\gamma_2,\theta=\theta_2\theta_1^{\gamma_2}\right] \end{aligned}$$

Because the four original parameters are positive, the two reduced parameters are defined and

---

[10] Consider the transformation $Y = X^{0.5}$, where $X^{0.5} \sim Gamma(½,2)$. Because the gamma distribution is medium-tailed and $\gamma = ½$, *Y* is light-tailed: $medium \times (\gamma < 1) = \infty = light$. Moreover:

$$f_Y(x) = \frac{1}{\Gamma(½)} e^{-\frac{x^2}{2}} \left(\frac{x^2}{2}\right)^{½-1} \frac{1}{2}\frac{dx^2}{dx} = \frac{1}{\sqrt{\pi}} e^{-\frac{x^2}{2}} \frac{\sqrt{2}}{x} x = \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = f_{|N(0,1)|}(x)$$

Hence, $Y \sim |N(0,1)|$; the negative support of the standard normal distribution has been reflected onto the positive support. Thus, the right tail of the normal distribution is light, in confirmation of what we derived in Section 3.

positive. By repetition, an *n*-step power transformation is always equivalent to a direct one. Due to the asymmetry of the scale formula, power transformation is not commutative. But since $\theta_3\left(\theta_2\theta_1^{\gamma_2}\right)^{\gamma_3} = \left(\theta_3\theta_2^{\gamma_3}\right)\theta_1^{\gamma_2\gamma_3}$, it is associative. And power transformation can always be inverted: $PT\left[PT[X;\gamma,\theta]; 1/\gamma, \theta^{-1/\gamma}\right] = PT[X;1,1] = X$.

So if *X* can power-transform into *Y*, it can do so in one step. And from *Y* it can return to *X* in one step. Therefore the range within which *X* can power-transform is a closed network.[11] The *power-transformation network* of *X* is the set $ptn(X) = \{Y \in \Xi : \exists\gamma,\theta > 0 : Y = PT[X;\gamma,\theta]\}$. And the *power-transformation range* of $\Phi \subseteq \Xi$, where Φ is a set of random variables (or of their distributions), can be defined as $PTR(\Phi) = \{Y \in \Xi : \exists X \in \Phi, \exists\gamma,\theta > 0 : Y = PT[X;\gamma,\theta]\}$ or as $PTR(\Phi) = \bigcup\{ptn(X) : X \in \Phi\}$. Unlike $ptn(X)$, there is no guarantee of a power-transformation connection between any two elements of $PTR(\Phi)$. Because 'network' connotes interconnectedness, we changed the noun here to 'range'.

Obviously, $PTR(\varnothing) = \varnothing$ and $PTR(\Xi) = \Xi$. But of interest here is $PTR(\Xi_+)$, the set of all distributions that can be formed by power-transforming medium-tailed distributions. Above we saw that the power transformation "knocks distributions off the medium ridge into light or heavy valleys." Therefore, this set is larger than Ξ_+, i.e., $\Xi_+ \subset PTR(\Xi_+)$. It spills into Ξ_0 and Ξ_∞, or in symbols $\Xi_0 \cap PTR(\Xi_+) \neq \varnothing$ and $\Xi_\infty \cap PTR(\Xi_+) \neq \varnothing$. But there are distributions in Ξ_0 and Ξ_∞ that are unattainable from Ξ_+ by power transformation. We found above that power transformation cannot unseat distributions that are so heavy as to have infinite moments; hence, a trip to them from Ξ_+ to them is precluded. But even the lognormal, whose moments are all finite, power-transforms back to lognormal.

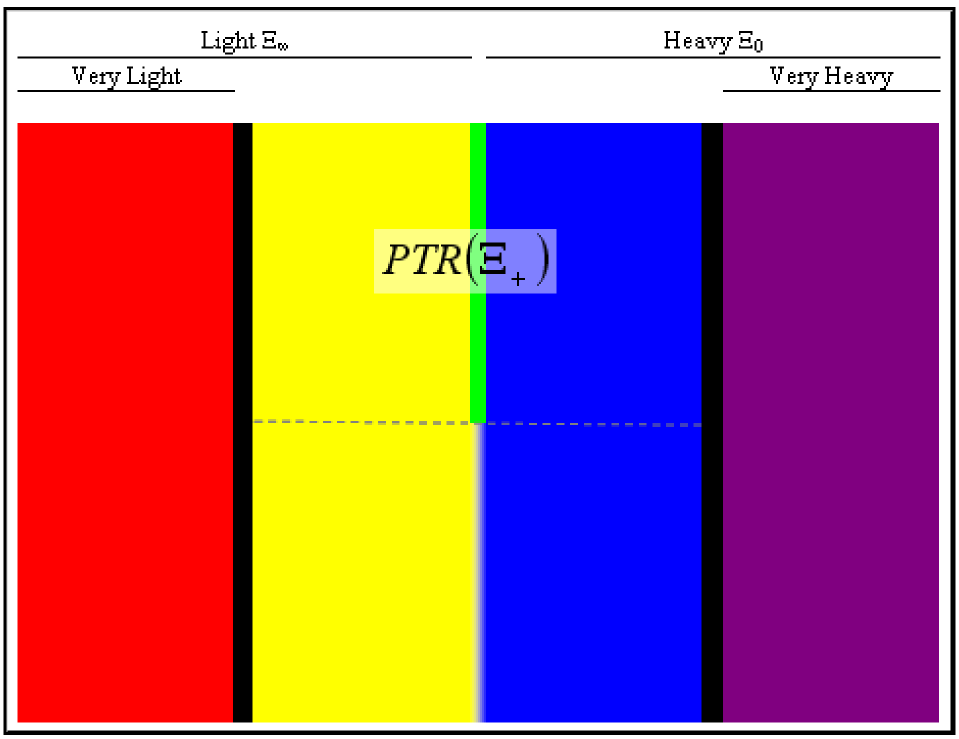
On the other hand, the symmetry of the multiplication table hints that some light-tailed distributions might be too light to power-transform elsewhere. Indeed, the survival function of one such distribution is $S_Q(x) = e^{-\left(e^x - 1\right)}$. It is light-tailed, since $\tau_Q = -\lim_{x\to\infty}\frac{d\,ln\,S_Q(x)}{dx} = \lim_{x\to\infty}\frac{d\left(e^x - 1\right)}{dx} = \infty$. But if $\frac{Y}{\theta} = Q^\gamma$, then:

---

[11] Technically, it is an algebraic group, whose set G is $\left\{\langle\gamma,\theta\rangle \in \Re^+ \times \Re^+\right\}$ and whose function $f : G \times G \to G$ is $f\left(\langle\gamma_1,\theta_1\rangle, \langle\gamma_2,\theta_2\rangle\right) = \left\langle\gamma_1\gamma_2, \theta_2\theta_1^{\gamma_2}\right\rangle$. The function is associative; $\langle 1, 1\rangle$ is the identity element; and every element has an inverse: $f\left(\langle\gamma,\theta\rangle, \left\langle 1/\gamma, \theta^{-1/\gamma}\right\rangle\right) = f\left(\left\langle 1/\gamma, \theta^{-1/\gamma}\right\rangle, \langle\gamma,\theta\rangle\right) = \langle 1, 1\rangle$.

$$\tau_Y = \frac{\gamma}{\theta}\lim_{v\to\infty}\left\{-\frac{d\,ln\,S_Q(v)}{dv}\cdot v^{\frac{\gamma-1}{\gamma}}\right\} = \frac{\gamma}{\theta}\lim_{v\to\infty}\left\{\frac{d\left(e^v-1\right)}{dv}\cdot v^{\frac{\gamma-1}{\gamma}}\right\} = \frac{\gamma}{\theta}\lim_{v\to\infty}\left\{e^v\cdot v^{\frac{\gamma-1}{\gamma}}\right\} = \infty\,.$$

Hence, this distribution remains light under power transformation. So within the $0\cdot\infty$ and $\infty\cdot 0$ cells of the table are distributions so heavy-tailed and so light-tailed as to be unmoved by power transformation. So these cannot belong to $PTR(\Xi_+)$. Because of this duality, we deem the power transformation to be a better basis for subclassification than the divergence of positive moments. Tail-class immutability to power transformation merits the adverb 'very'. Thus we will now speak of "very light-tailed" and "very heavy-tailed" distributions and random variables. The lognormal is very heavy-tailed, though not as heavy-tailed as something with missing moments. Quite appropriately, nothing is "very" medium-tailed; medium is just medium. The following diagram will conclude this section:



The diagram, which looks like a painted tennis court with half a net, represents a tripartite form of Ξ. The black regions are boundaries; Ξ is the union of the colored regions. The middle partition is the set of all loss distributions whose tail classes change under power transformation. All the medium-tailed distributions, the green area, must belong to this set. The red and violet regions contain all the loss distributions whose tail classes do not change. These unchangeable distributions are either light-tailed and in the red region or heavy-tailed and in the violet region. The yellow region contains the changeable light-tailed distributions, the blue the changeable heavy-tailed. By definition, power transformation cannot "jump" from the red or violet regions. But if perchance, it could jump from the middle, it could not jump back. So since power transformation is reversible, the black regions are barriers to power transformation.

Now consider all the changeable distributions as organized into horizontal slices of power-transformation networks. Whatever might be the position of distribution X in its network, $Y = X^{\gamma}$ transforms X to the right if $\gamma > 1$ and to the left if $\gamma < 1$. The movement approaches the black boundaries as γ approaches infinity or zero. If the movement passes through a medium-tailed distribution, then it is within the power-transform range $PTR(\Xi_{+})$. Since a power-transformation range can contain at most one medium-tailed distribution,[12] the set of medium-tailed distributions Ξ_{+} is merely an interface between Ξ_0 and Ξ_{∞}. It is not intended for the green region to appear thick.

But one might think that the transition between light and heavy implies that some X^{γ} is medium-tailed. A counterexample disproves this: let R be the random variable $S_R(x) = (1+x)^{-x}$. Its hazard rate is $\lambda_R(x) = ln(1+x) + x/(1+x)$. Hence:

$$
\begin{aligned}
\tau_{\frac{Y}{\theta}=R^{\gamma}} &= \frac{1}{\gamma\theta} \lim_{v\to\infty} \left\{ \lambda_R(v) \cdot v^{1-\gamma} \right\} \\
&= \frac{1}{\gamma\theta} \lim_{v\to\infty} \left\{ \left(ln(1+v) + v/(1+v) \right) \cdot v^{1-\gamma} \right\} \\
&= \frac{1}{\gamma\theta} \lim_{v\to\infty} \left(\frac{ln(1+v) + v/(1+v)}{ln(v)} \cdot ln(v) \cdot v^{1-\gamma} \right) \\
&= \frac{1}{\gamma\theta} \lim_{v\to\infty} \left(ln(v) \cdot v^{1-\gamma} \right) \\
&= \begin{cases} \infty & 0 < \gamma \le 1 \\ 0 & \gamma > 1 \end{cases}
\end{aligned}
$$

So there are power transformations back and forth between Ξ_{∞} and Ξ_0 which avoid Ξ_{+}. For this reason, the area underneath the green is porous; it shades from yellow to blue

[12] More accurately, it contains at most one medium-tailed distribution per θ. The diagram cannot represent Ξ as a metric space; only τ_X and X^{γ} are represented.

8. THE ULTIMATE SETTLEMENT RATE UNDER EXPONENTIAL AND LOGARITHMIC TRANSFORMATIONS

Our third transformation is the exponential, which Klugman defines as $Y = e^X$ [4, p. 95]. However, for two reasons we prefer the form $\frac{Y}{\theta} = \frac{e^{\eta X} - 1}{\eta}$, for $\eta > 0$.[13] First, although $Y = e^X$ works for such random variables as the normal, with support over $\Re$, we are transforming loss distributions, whose support is positive. We wish all our transformations $y = h(x)$ to be strictly increasing functions from $[0, \infty)$ *onto*, not just *into*, $[0, \infty)$. Therefore, $0 = h(0)$, as it does in the above forms. Second, the parameter η, though not strictly necessary, standardizes the transformation; it sets the derivative at zero to unity, or $1 = h'(0)$. The standardized transformation looks like $y = x$ in the neighborhood of the origin. In fact, the limit of the standardized transformation as $\eta \to 0^+$ is the identity function $h(x) = x$. The appeal of this limiting case is the second reason for our form.

As for the ultimate settlement rate under the exponential transformation:

$$S_Y(u) = Prob\left[\frac{Y}{\theta} > \frac{u}{\theta}\right] = Prob\left[\frac{e^{\eta X} - 1}{\eta} > u/\theta\right] = Prob\left[X > \frac{1}{\eta} ln\left(1 + \eta\frac{u}{\theta}\right)\right] = S_X\left(\frac{1}{\eta} ln\left(1 + \eta\frac{u}{\theta}\right)\right)$$

Therefore:

---

[13] Actually, for precision and to ensure unitless parameters in transcendental functions we should include two scale parameters: $\frac{Y}{\theta} = \frac{e^{\eta\frac{X}{\phi}} - 1}{\eta}$. But, again, as in Footnote 5, this overparameterizes, for $\frac{Y}{(\theta/\phi)} = \frac{e^{(\eta/\phi)X} - 1}{(\eta/\phi)}$. This transformation is valid and meaningful even for $\eta < 0$, as explained in the appendix.

$$\begin{aligned}
\tau_Y &= -\lim_{u\to\infty}\frac{d\,ln\,S_Y(u)}{du} \\
&= -\lim_{u\to\infty}\left\{\frac{d\,ln\,S_X\left(\frac{1}{\eta}ln\left(1+\eta\frac{u}{\theta}\right)\right)}{d\left(\frac{1}{\eta}ln\left(1+\eta\frac{u}{\theta}\right)\right)}\frac{d\left(\frac{1}{\eta}ln\left(1+\eta\frac{u}{\theta}\right)\right)}{du}\right\} \\
&= \lim_{\substack{u\to\infty \\ v(u)\to\infty}}\left\{-\frac{d\,ln\,S_X(v)}{dv}\cdot\frac{1}{1+\eta\frac{u}{\theta}}\cdot\frac{1}{\theta}\right\} \\
&= \lim_{v\to\infty}\left\{\lambda_X(v)\cdot\frac{1}{e^{\eta v}}\cdot\frac{1}{\theta}\right\} \\
&= \frac{1}{\theta}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot e^{-\eta v}\right\} \\
&= \frac{1}{\theta}\tau_X\cdot 0
\end{aligned}$$

Since $\eta > 0$, medium- and heavy-tailed random variables exponentially transform into heavy-tailed ones; light-tailed random variables are indeterminate.

But let *X* be light-tailed ($\tau_X = \infty$), but not very light-tailed. This puts *X* in the yellow region of the diagram. Using "simply" for "not very," we can say that *X* is simply light-tailed. Then, it becomes heavy-tailed ($\tau_Z = 0$) under some $\gamma > 1$ power transformation $Z = X^\gamma$. Hence:

$$0 = \gamma\cdot 0 = \gamma\cdot\tau_Z = \gamma\cdot\frac{1}{\gamma}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot v^{1-\gamma}\right\} = \lim_{v\to\infty}\left\{\lambda_X(v)\cdot v^{1-\gamma}\right\}$$

This information resolves the indeterminacy of the exponential transformation. The following proof makes use of the truth that $\lim_{v\to\infty}\left\{e^{-\eta v}v^{\gamma-1}\right\} = 0$ for $\eta > 0$:

$$\begin{aligned}0 &= \frac{1}{\theta}\cdot 0\cdot 0\\ &= \frac{1}{\theta}\cdot \lim_{v\to\infty}\left\{\lambda_X(v)\cdot v^{1-\gamma}\right\}\cdot \lim_{v\to\infty}\left\{e^{-\eta v}v^{\gamma-1}\right\}\\ &= \frac{1}{\theta}\cdot \lim_{v\to\infty}\left\{\lambda_X(v)\cdot v^{1-\gamma}e^{-\eta v}v^{\gamma-1}\right\}\\ &= \frac{1}{\theta}\cdot \lim_{v\to\infty}\left\{\lambda_X(v)\cdot e^{-\eta v}\right\}\\ &= \tau_Y\end{aligned}$$

So, in the "simply light" case, the indeterminacy of $\tau_Y = \frac{1}{\theta}\tau_X \cdot 0 = \frac{1}{\theta}\cdot\infty\cdot 0$ resolves to heavy.

However, we do not yet know whether *Y* is simply heavy or very heavy (blue or violet). So now let *Z* now be a power transformation of $\frac{Y}{\theta}$, i.e., $Z = \left(\frac{Y}{\theta}\right)^{\gamma} = \left(\frac{e^{\eta X}-1}{\eta}\right)^{\gamma}$. And so:

$$S_Z(u) = Prob\left[\left(\frac{e^{\eta X}-1}{\eta}\right)^{\gamma} > u\right] = Prob\left[X > \frac{1}{\eta}ln\left(1+\eta u^{\frac{1}{\gamma}}\right)\right] = S_X\left(\frac{1}{\eta}ln\left(1+\eta u^{\frac{1}{\gamma}}\right)\right)$$

We have seen just above that because *X* is simply light-tailed, $\lim_{v\to\infty}\left\{\lambda_X(v)\cdot e^{-\eta v}\right\} = 0$. But this holds true any $\eta > 0$. And since $\gamma > 0$, it will hold true also for $\gamma\eta > 0$. Therefore, knowing that $\lim_{v\to\infty}\left\{\lambda_X(v)\cdot e^{-\gamma\eta v}\right\} = 0$, we can determine the value of τ_Z:

$$\tau_Z = -\lim_{u\to\infty}\left\{\frac{d\,ln\,S_X\left(\frac{1}{\eta}ln\left(1+\eta u^{\frac{1}{\gamma}}\right)\right)}{d\left(\frac{1}{\eta}ln\left(1+\eta u^{\frac{1}{\gamma}}\right)\right)}\frac{d\left(\frac{1}{\eta}ln\left(1+\eta u^{\frac{1}{\gamma}}\right)\right)}{du}\right\}$$

$$= \lim_{\substack{u\to\infty\\ v(u)\to\infty}}\left\{-\frac{d\,ln\,S_X(v)}{dv}\cdot\frac{\frac{1}{\gamma}u^{\frac{1}{\gamma}-1}}{1+\eta u^{\frac{1}{\gamma}}}\right\}$$

$$= \lim_{v\to\infty}\left\{\lambda_X(v)\cdot\frac{\frac{1}{\gamma}\left(\frac{e^{\eta v}-1}{\eta}\right)^{1-\gamma}}{e^{\eta v}}\right\}$$

$$= \frac{1}{\gamma\cdot\eta^{1-\gamma}}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot\left(\frac{e^{\eta v}-1}{e^{\eta v}}\right)^{1-\gamma}\cdot\frac{1}{e^{\gamma\eta v}}\right\}$$

$$= \frac{1}{\gamma\cdot\eta^{1-\gamma}}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot e^{-\gamma\eta v}\right\}$$

$$= \frac{1}{\gamma\cdot\eta^{1-\gamma}}\cdot 0$$

$$= 0$$

Consequently, Y is a heavy-tailed random variable whose tail class is invariant to power transformation; it is very heavy-tailed. This proves that an exponential transformation of a simply light-tailed random variable is a very heavy-tailed random variable. In the diagram exponential transformation moves from the yellow region to the violet; unlike power transformation, it is capable of jumping a least the right barrier.

In the exponential-transformation of medium- and heavy-tailed random variables, there is no indeterminacy to $\tau_Y = \frac{1}{\theta}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot e^{-\eta v}\right\} = 0$. But again, the ultimate settlement rate of a subsequent power transformation is $\tau_Z = \frac{1}{\gamma\cdot\eta^{1-\gamma}}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot e^{-\gamma\eta v}\right\} = 0$. Hence, exponential transformation turns medium and simply heavy tails into very heavy tails. In sum, it transforms yellow, green, and blue into violet. But its differing effect on moments will be treated in the next section.

We can be brief about the logarithmic transformation $\frac{Y}{\theta} = \frac{1}{\eta} ln(1+\eta X)$. It is the inverse of the exponential transformation, but less obviously than in the case of power transformation. Define two operators: the exponential transformation $ET[X;\eta,\theta] = \theta\frac{e^{\eta X}-1}{\eta}$ and the logarithmic transformation $LT[X;\eta,\theta] = \theta\cdot\frac{1}{\eta} ln(1+\eta X)$. Just as $ET[X;\eta,\theta] \to \theta X$ as $\eta \to 0^+$, so too $LT[X;\eta,\theta] \to \theta X$ as $\eta \to 0^+$. One who performs the algebra will find that $ET\left[LT[X;\eta,\theta];\frac{\eta}{\theta},\frac{1}{\theta}\right] = LT\left[ET[X;\eta,\theta];\frac{\eta}{\theta},\frac{1}{\theta}\right] = X$. Since $\eta, \theta > 0$, the inverting parameters exist. Now if $\frac{Y}{\theta} = \frac{1}{\eta} ln(1+\eta X)$, then $S_Y(u) = S_X\left(\frac{e^{\frac{\eta u}{\theta}}-1}{\eta}\right)$. The reader should be able to prove by now that $\tau_Y = \frac{1}{\theta}\lim_{v\to\infty}\{\lambda_X(v)\cdot(1+\eta v)\} = \frac{1}{\theta}\tau_X\cdot\infty$. So $\eta > 0$ logarithmic transformation turns light and medium into light; the heavy-tailed random variables are indeterminate. It is not necessary to repeat the power-on-top-of-exponential-transformation argument. Because of exponential-logarithmic inversion, the question "Into what do yellow, green, and blue logarithmically transform?" is equivalent to "What exponentially transforms into yellow, green, and blue?" Whatever it is, it can't be undone by power transformation. Therefore, what exponentially transforms into the middle region of the diagram must be very light. So exponential and logarithmic transformations from the middle jump the barriers.

9. POSITIVE MOMENTS AND THE EXPONENTIAL TRANSFORMATION

Section 5 proved that all the positive moments of light- and medium-tailed random variables are finite. An infinite moment is a sufficient, but not a necessary, condition for a heavy tail. Here we will examine the positive moments of the exponentially transformed $Y = \frac{e^{\eta X}-1}{\eta}$. But $Y^k = \left(\frac{e^{\eta X}-1}{\eta}\right)^k$. Although this is on the order of $e^{k\eta X}$, it is not the same. Since our findings depend on the behavior of $e^{k\eta X}$, we must first prove that $E[Y^k]$ is finite if and only if $E[e^{k\eta X}]$ is finite.

To begin, $E[Y^k] = E\left[\left(\frac{e^{\eta X}-1}{\eta}\right)^k\right] = \frac{E\left[\left(e^{\eta X}-1\right)^k\right]}{\eta^k}$. Since η and k are positive, η^k is positive. So $E[Y^k]$ is finite if and only $E\left[\left(e^{\eta X}-1\right)^k\right]$ is finite. And since $0 \le e^{\eta x}-1 < e^{\eta x}$ over the support of X, $0^k = 0 \le \left(e^{\eta x}-1\right)^k < e^{k\eta x}$. So $Prob\left[0 \le \left(e^{\eta X}-1\right)^k < e^{k\eta X}\right] = 1$ and $0 \le E\left[\left(e^{\eta X}-1\right)^k\right] < E\left[e^{k\eta X}\right]$. Therefore, if $E\left[e^{k\eta X}\right]$ is finite, then so too is $E\left[\left(e^{\eta X}-1\right)^k\right]$. As for the converse:

$$
\begin{aligned}
E\left[\left(e^{\eta X}-1\right)^k\right] &= \int_{x=0}^{\infty}\left(e^{\eta x}-1\right)^k dF_X(x) \\
&= \int_{\eta x=0}^{\infty}\left(e^{\eta x}-1\right)^k dF_X(x) \\
&= \int_{\eta x=0}^{ln\,2}\left(e^{\eta x}-1\right)^k dF_X(x)+\int_{\eta x=ln\,2}^{\infty}\left(e^{\eta x}-1\right)^k dF_X(x) \\
&= \int_{\eta x=0}^{ln\,2}\left(e^{\eta x}-1\right)^k dF_X(x)+\int_{\eta x=ln\,2}^{\infty}\left(\frac{e^{\eta x}-1}{e^{\eta x}}\right)^k e^{k\eta x} dF_X(x) \\
&\geq \int_{\eta x=0}^{ln\,2}(0)^k dF_X(x)+\int_{\eta x=ln\,2}^{\infty}\left(\frac{1}{2}\right)^k e^{k\eta x} dF_X(x) \\
&\geq \left(\frac{1}{2}\right)^k \int_{\eta x=ln\,2}^{\infty} e^{k\eta x} dF_X(x)
\end{aligned}
$$

Consequently, $\int_{\eta x=ln\,2}^{\infty} e^{k\eta x} dF_X(x) \leq 2^k\, E\left[\left(e^{\eta X}-1\right)^k\right]$. Furthermore:

$$
\begin{aligned}
E\left[e^{k\eta X}\right] &= \int_{\eta x=0}^{\infty} e^{k\eta x} dF_X(x) \\
&= \int_{\eta x=0}^{ln\,2} e^{k\eta x} dF_X(x)+\int_{\eta x=ln\,2}^{\infty} e^{k\eta x} dF_X(x) \\
&\leq \int_{\eta x=0}^{ln\,2} 2^k\, dF_X(x)+\int_{\eta x=ln\,2}^{\infty} e^{k\eta x} dF_X(x) \\
&\leq 2^k + \int_{\eta x=ln\,2}^{\infty} e^{k\eta x} dF_X(x) \\
&\leq 2^k + 2^k\, E\left[\left(e^{\eta X}-1\right)^k\right]
\end{aligned}
$$

So $0 \leq E\left[e^{k\eta X}\right] \leq 2^k + 2^k\, E\left[\left(e^{\eta X}-1\right)^k\right]$. Therefore, if $E\left[\left(e^{\eta X}-1\right)^k\right]$ is finite, so too is $E\left[e^{k\eta X}\right]$. Thus have we shown that $E\left[Y^k\right]$ is finite if and only if $E\left[e^{k\eta X}\right]$ is finite.

Now we continue with the simpler problem of examining the moments of $E\left[e^{k\eta X}\right]$. Using again the theorem from Section 5 that $E[h(X)] = h(0) + \int_{x=0}^{\infty} S_X(x)dh(x)$, we have:

$$
E\left[e^{k\eta X}\right] = e^{k\eta 0} + \int_{x=0}^{\infty} S_X(x)de^{k\eta x} = 1 + k\eta \int_{x=0}^{\infty} S_X(x)e^{k\eta x}dx
$$

Therefore, $E\left[Y^k\right]$ is finite if and only if $\int_{x=0}^{\infty} S_X(x)e^{k\eta x}dx$ is finite. Dispensing with mathematical rigor, we know that $\int_{x=0}^{\infty} S_X(x)e^{k\eta x}dx$ is finite if and only there exists a $\xi > 0$ such that for all $x \geq \xi$, $S_X(x) \leq S_X(\xi)e^{-k\eta(x-\xi)}$. In words, in order for the integral to converge, at some point the survival function must decay at a rate greater than $k\eta$. But the limit of this decay is the ultimate settlement rate τ_X. Hence, $E\left[Y^k\right]$ is finite if and only if $k\eta < \tau_X$, or $k < \tau_X/\eta$. Therefore, all the positive moments of exponential transformations of light-tailed ($\tau_X = \infty$) distributions are finite. The positive moments of exponential transformations of medium-tailed ($0 < \tau_X < \infty$) distributions are finite for $k < \tau_X/\eta$ and infinite for $k \geq \tau_X/\eta$. And all the positive moments of exponential transformations of heavy-tailed ($\tau_X = 0$) distributions are infinite.[14]

10. INVERTING AND MIXING LOSS DISTRIBUTIONS

The commentary on property (i) in Section 2 stated the desirability for a loss distribution to be invertible. But our only use of an inverse distribution was to derive the ultimate settlement rate of the inverse gamma in Section 3. Moreover, Klugman lists a fourth transformation, viz., mixing [4, pp. 97-99]. Both inverting and mixing are involved in the generalized Pareto, because:

$$\begin{aligned} GenPareto(\alpha,\beta,\theta) &\sim \frac{\theta}{Gamma(\beta,1)} \cdot Gamma(\alpha,1) \\ &\sim InvGamma(\beta,\theta) \cdot Gamma(\alpha,1) \\ &\sim Gamma(\alpha, InvGamma(\beta,\theta)) \end{aligned}$$

So the generalized Pareto can be formed as a gamma distribution whose scale parameter is an inverse-gamma distribution. In this section we will explain why inverting and mixing tend to produce heavy-tailed distributions.

[14] Thus indirectly we verify what we know about the lognormal distribution, an exponential transformation of the light-tailed normal, viz., that all its moments are finite: $E\left[e^{kN(\mu,\sigma^2)}\right] = e^{k\mu + k^2\sigma^2/2}$. $E\left[Y^k\right] \approx E\left[e^{k\eta X}\right] = M_X(k\eta)$. If $X \sim Gamma(\alpha,\theta)$, $M_X(k\eta) = (1-\theta k\eta)^{-\alpha}$, which diverges for $k\eta \geq 1/\theta = \tau_X$. The missing moments explain the intractability, or even the nonexistence, of the moment generating functions of all but the simplest distributions. However, the zeroth moment is finite: $E\left[Y^0\right] = E\left[e^{0X}\right] = 1$. So by virtue of absolute convergence in the complex numbers, viz., $|\varphi_X(t)| = \left|E\left[e^{itX}\right]\right| \leq E\left[\left|e^{itX}\right|\right] = E[1] = 1$, the characteristic function is more successful. All the imaginary moments of all real random variables exist as complex numbers.

First, as to inverting, if $Y = 1/X$, then:

$$S_Y(u) = Prob\left[\frac{1}{X} > u\right] = Prob\left[X < \frac{1}{u}\right] = 1 - Prob\left[X \geq \frac{1}{u}\right] = 1 - Prob\left[X = \frac{1}{u}\right] - S_X\left(\frac{1}{u}\right)$$

Here we will assume that *X* has no probability mass, at least not in the neighborhood of zero. So $S_Y(u) = 1 - S_X\left(\frac{1}{u}\right)$. Therefore:

$$\lambda_Y(u) = -\frac{d\,ln\,S_X(x)}{dx} = -\frac{d\,ln\left(1 - S_X\left(\frac{1}{u}\right)\right)}{dx} = \frac{-S'_X\left(\frac{1}{u}\right)}{1 - S_X\left(\frac{1}{u}\right)}\frac{1}{u^2} = \frac{f_X\left(\frac{1}{u}\right)}{1 - S_X\left(\frac{1}{u}\right)}\frac{1}{u^2}$$

And so the ultimate settlement rate of *Y* is:

$$\tau_Y = \lim_{u\to\infty}\lambda_Y(u) = \lim_{u\to\infty}\frac{f_X\left(\frac{1}{u}\right)}{1 - S_X\left(\frac{1}{u}\right)}\frac{1}{u^2} = \lim_{v\to 0^+}\frac{f_X(v)}{1 - S_X(v)}v^2 = \lim_{v\to 0^+}\frac{f_X(v)}{F_X(v)}v^2 = \lim_{v\to 0^+}\frac{f_X(v)}{h(v)}v,$$

where $h(v) = \frac{F_X(v)}{v} = \frac{1}{v}\int_{x=0}^{v} f_X(x)$. It is the average height of f_X over the interval $(0, v]$. If f_X increases as $x \to 0^+$, $h(v) > f_X(v) > 0$. Then $0 < \frac{f_X(v)}{h(v)} < 1$, and $\tau_Y = 0$. This holds true even if f_X approaches infinity. If f_X is bounded within two positive numbers, then again, $\tau_Y = 0$. The remaining possibility is that f_X decreases to zero as $x \to 0^+$, in which case $1 < \frac{f_X(v)}{h(v)}$.

Now if f_X is zero in some interval $[0, \varepsilon]$, then $Prob[X \leq \varepsilon] = 0$; so $S_Y(1/\varepsilon) = Prob[Y > 1/\varepsilon] = 0$ and property (ii) would disqualify *Y* as a loss distribution. So f_X decreases to zero, but equals zero only at the origin. The obvious choice is a power-function approach into the origin, i.e., $f_X(v) \propto v^{\gamma-1}$ for $\gamma > 1$. But then $F(v) \propto v^{\gamma}/\gamma$ and $\tau_Y = \lim_{v\to 0^+}\frac{v^{\gamma-1}}{\left(v^{\gamma}/\gamma\right)}v^2 = 0$. So even power-function approaches are not slow enough. For τ_Y to be positive, near the origin, $\frac{f_X(v)}{F_X(v)}$ must be on the order of v^{-2}. So $\frac{d\,ln\,F_X(v)}{dv} = \frac{f_X(v)}{F_X(v)} = \frac{k}{v^2}$, and $ln\,F_X(v) - ln\,F_X(\varepsilon) = \int_{x=\varepsilon}^{v}\frac{k}{x^2}dx = \frac{k}{\varepsilon} - \frac{k}{v}$. Thus, $F_X(v) = F_X(\varepsilon)e^{\frac{k}{\varepsilon}}e^{-\frac{k}{v}}$. The solution which satisfies $F_X(0) = 0$ is $F_X(x) = \begin{cases} e^{-\frac{k}{x}} & x > 0 \\ 0 & x = 0 \end{cases}$. But this is the inverse exponential cumulative density function. So, the only likely way of obtaining anything other than a heavy-tailed distribution by inversion is to invert an already inverted distribution. One may expect inversion to produce heavy-tailed distributions.

As for mixing, let $S_X(x) = \int_{\theta} S_{X|\theta}(x)dh(\theta)$. The survival function of the mixed distribution is the weighting according to $dh(\theta)$ of the distributions indexed by θ. Hence:

$$\lambda_X(x) = -\frac{S'_X(x)}{S_X(x)}$$

$$= \frac{\int_\theta S'_{X|\theta}(x)dh(\theta)}{\int_\theta S_{X|\theta}(x)dh(\theta)}$$

$$= \frac{\int_\theta \lambda_{X|\theta}(x)\cdot S_{X|\theta}(x)dh(\theta)}{\int_\theta S_{X|\theta}(x)dh(\theta)}$$

$$= \int_\theta \lambda_{X|\theta}(x)\cdot dw(x,\theta)$$

In the last equation $dw(x,\theta) = S_{X|\theta}(x)dh(\theta) \Big/ \int_\theta S_{X|\theta}(x)dh(\theta)$. The weights vary by *x*; but for all *x*, $\int_\theta dw(x,\theta) = 1$. As $x \to \infty$, the weighting will shift more and more toward the "surviving" distributions, i.e., in favor of the distributions whose τ is least. Consequently, $\tau_X = \inf\{\tau_{X|\theta}\}$. For the mixed exponential distribution $S_{MX}(x) = \sum_{i=1}^{n} p_i e^{-\frac{x}{\theta_i}}$, $\tau_X = min(1/\theta_i) = 1/max(\theta_i)$. This is medium-tailed; but $S_{MX}(x) = \sum_{i=1}^{\infty} p_i e^{-\frac{x}{\theta_i}}$ is heavy-tailed, if $\lim_{i\to\infty} \theta_i = \infty$.[15]

11. SUPER LIGHT AND SUPER HEAVY

The region within the barriers of the diagram is like the everyday world. Its span is that of the power transformation. But of course, in the long run e^x overwhelms x^n. To what others mean loosely by "in the long run"[16] mathematicians have given precision, viz., $\lim_{x\to\infty}$.[17] Just over the right

---

[15] The Tweedie distribution is $T = X_1 + \ldots + X_N$, for $X \sim Gamma(\alpha, \theta)$ and $N \sim Poisson(\lambda)$. Therefore, $T|N \sim Gamma(N\alpha, \theta)$ and $\tau_{T|N} = 1/\theta$. So, $\tau_T = \inf\{\tau_{T|N}\} = 1/\theta$, and T is medium-tailed.

[16] Such loose speech harbors specious arguments, for which Keynes expressed disdain in his famous quip, "In the long run we're all dead." Many use the adverb 'exponentially', as in "Something is growing exponentially," to express alarm, as if dealing with that thing were a critical matter. The sober truth is that almost all growth is exponential, but of limited duration. Mathematically, for $x \approx 0$, $e^{\gamma x} \approx 1 + \gamma x$. Moreover, the obverse is never considered: no one ever expresses alarm by saying that something is decaying exponentially.

[17] Still amazing even after 150 years are the accomplishments of such mathematicians as Cauchy, Weierstrass, and Dedekind concerning the nature of the real numbers, which finally put to rest the 2300-year-old paradoxes of Zeno. One who might try to resurrect them on the basis of today's quantum theory would ignore the fact that the paradoxes themselves presuppose continuity.

barrier are some familiar enough distributions, the very heavy-tailed ones. Hopping over it, first we'll find exponential transformations of simply light tails with all their moments. Second we'll find ETs of middle tails with moments up to a point. Third we'll find ETs of simply heavy tails with no moments at all. The familiar enough distributions are of the first two types; the distributions with no moments we will call "super heavy-tailed." We could refine our diagram's color scheme in conformity with the spectrum: first indigo, second violet, and third ultraviolet. The span of these distributions is on the order of $e^{\eta\left(x^{\gamma}\right)}$. But because $e^{\left(x^{\gamma}\right)} >> \left(e^{x}\right)^{\gamma}$, their span is greater than that of the power transformation. In fact, their span is "power-on-top-of-exponential." But at the end of that span is another barrier, over which another exponential transformation jumps, and so forth. The same applies in the other direction, into the microworld, with the logarithmic transformation. In descending order of heaviness are logarithmic transforms of simply heavy tails, which we could color orange. Second are LTs of middle tails, which remain red. And third are LTs of simply light tails, "super light-tailed" distributions whose color is infrared. And then we find a barrier to be surmounted by another LT.[18] So the classification is indefinitely extendable; but current needs remain within one transformation of the center.

11. CONCLUSION

Good classifications are not arbitrary; they are not set by convention or decree. Natural classifications should actually help those who study a subject to understand it and eventually to make deeper discoveries. Work is made easier with the right tools, and the essential tool for intellectual work is clear definition and classification. In this paper we entered the house of loss distributions through the door of the medium-tail distribution. We explored the first floor with the help of the power transformation, and then found exponential and logarithmic staircases to the second floor and the basement. Some the mathematics was formidable; but it all reduces to the interaction between power and exponential functions. The classification scheme yielded new and beautiful insights. Surely there is much more to be discovered; but the classification of distributions into light, medium, and heavy, as well as the subclassifications "very" and "super," almost as surely will play an important role therein. Though it might be hard for now to put this theory to practical use (we've given no list of "which distribution for which purpose"), actuaries have a right to appreciate the beauty of their subject – its æsthetic value. And many, perhaps most, practical benefits have arisen from what once had been considered "mere theory."[19]

---

[18] But however rarefied these tails may become, they are still infinite.

[19] That good theory aids discovery and technological progress (and conversely, that bad theory impedes them) is illustrated in modern physics. On the basis of quantum theory in 1928 Wolfgang Pauli predicted the existence of antimatter, in particular, the anti-electron or positron, which was discovered in 1932 and whose discovery now benefits mankind in positron emission tomography – commonly performed in hospitals as PET scans. Since the 1960s the standard model of particle physics has predicted the one still missing particle, the Higgs boson, whose existence many

REFERENCES

[1] Corro, Dan, "Grouping Loss Distributions by Tail Behavior," *CAS E-Forum* (Fall 2008), 448-541, www.casact.org/pubs/forum/08fforum.

[2] Embrechts, Paul, Resnick, Sidney I., and Samorodnitsky, Gennady, "Extreme Value Theory as a Risk Management Tool," *North American Actuarial Journal*, 3:2, 1999, 30-41, www.casact.org/library/studynotes/Embrechts_ExtremeValue.pdf.

[3] Halliwell, Leigh J., "The Mathematics of Excess Losses," *Variance*, 6:1, 2012, 32-47, www.variancejournal.org/issues/06-01/32.pdf.

[4] Klugman, Stuart A., and Panjer, Harry H., and Willmot, Gordon E., *Loss Models: From Data to Decisions*, 1998, John Wiley & Sons, New York.

[5] Meyers, Glenn, "Pure Premium Regression with the Tweedie Model," The Actuarial Review (May 2009), 12-13, www.casact.org/newsletter/pdfUpload/ar/Brainstorms_corrected_from AR_May2009.pdf.

[6] Wikipedia contributors, "Generalized Pareto distribution," *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Generalized_Pareto_distribution.

---

physicists believe is about to be verified (as of January 2013) in CERN's Large Hadron Collider. Although science is inductive, it is not haphazard. On the basis of theory scientists make hunches. Theory is essential for posing the right questions, and helps in determining how and where to look for their answers. The present author discovered much about loss distributions while developing this classification, the following three things in particular. First, it led to his discovery of super heavy-tailed distributions; before he would have insisted that every distribution must have some positive moments. Second, the symmetry of the "multiplication table" of Section 6 beckoned the existence of super light-tailed distributions, one of which he then readily found, the *Q* distribution of Section 7. Third, again in Section 7, he spent considerable time trying to prove that every power-transformation range must include a medium-tailed distribution, i.e., that the green line of the diagram cuts all the way through. Twice he thought to have proven it only on checking to be disappointed. This opened him to try to falsify the proposition, during which he discovered a new expression for the ultimate settlement rate that specified a sufficient characteristic for a counterexample, which again he readily found, the *R* distribution. Thus, he is convinced that this classification is not arbitrary, no mere convention or convenience. Rather, it is fertile of discovery.

APPENDIX

Extreme Value Theory

Most casualty actuaries have studied the forms of loss distributions that are given in Klugman [3, Appendix A]. However, in the field of extreme-value theory, there is a generalized-Pareto distribution that differs from our $GenPareto(\alpha,\beta,\theta)$. In this appendix we will translate it into forms more familiar to actuaries. The survival function of this generalized-Pareto is:

$$S_X(x \geq 0;\xi,\theta) = Prob[X > x] = \begin{cases} \left(1+\xi\frac{x}{\theta}\right)^{-\frac{1}{\xi}} & \text{for } \xi \neq 0 \\ e^{-\frac{x}{\theta}} & \text{for } \xi = 0 \end{cases}$$

This is the definition given in [2, p. 33] and [5], except that we have zeroed a location parameter and used 'θ' instead of 'σ' for the scale parameter. The shape parameter ξ may be any real number, but the scale parameter θ, as always, must be positive. The function is defined for $\xi = 0$ as $S_X(x;0,\theta) = \lim_{\xi \to 0} S_X(x;\xi,\theta)$, which pertains to the $Gamma(1,\theta)$ or $Exponential(\theta)$ distribution.

For $\xi \neq 0$, the probability density function is $f_X(x) = -\frac{dS_X(x)}{dx} = \frac{1}{\xi}\left(1+\xi\frac{x}{\theta}\right)^{-\frac{1}{\xi}-1}\frac{\xi}{\theta}$. If $\xi > 0$, the exponent is negative. In this case, the function translates as follows:

$$f_X(x) = \frac{1}{\xi}\left(1+\xi\frac{x}{\theta}\right)^{-\frac{1}{\xi}-1}\frac{\xi}{\theta}$$

$$= \frac{1}{\xi}\left(\frac{1}{x/(\theta/\xi)+1}\right)^{\frac{1}{\xi}+1}\frac{1}{(\theta/\xi)}$$

$$= \frac{1}{\xi}\left(\frac{x/(\theta/\xi)}{x/(\theta/\xi)+1}\right)^{1-1}\left(\frac{1}{x/(\theta/\xi)+1}\right)^{\frac{1}{\xi}-1}\frac{1/(\theta/\xi)}{(x/(\theta/\xi)+1)^2}$$

$$= \frac{\Gamma\left(1+\frac{1}{\xi}\right)}{\Gamma(1)\Gamma\left(\frac{1}{\xi}\right)}\left(\frac{x/(\theta/\xi)}{x/(\theta/\xi)+1}\right)^{1-1}\left(\frac{1}{x/(\theta/\xi)+1}\right)^{\frac{1}{\xi}-1}\frac{1/(\theta/\xi)}{(x/(\theta/\xi)+1)^2}$$

$$= f_{GenPareto\left(1,\frac{1}{\xi},\frac{\theta}{\xi}\right)}(x) \quad \Rightarrow \quad \frac{X}{(\theta/\xi)} \sim \frac{Gamma(1,1)}{Gamma\left(\frac{1}{\xi},1\right)}$$

If $\xi < 0$, $-\frac{1}{\xi}$ is positive, and the translation is:

$$f_X(x) = \frac{1}{-\xi}\left(1-\xi\frac{x}{\theta}\right)^{\frac{1}{-\xi}-1}\frac{-\xi}{\theta}$$

$$= \frac{1}{-\xi}(1-x/(\theta/-\xi))^{\frac{1}{-\xi}-1}(1/(\theta/-\xi))$$

$$= \frac{1}{-\xi}(x/(\theta/-\xi))^{1-1}(1-x/(\theta/-\xi))^{\frac{1}{-\xi}-1}(1/(\theta/-\xi))$$

$$= \frac{\Gamma(1+1/-\xi)}{\Gamma(1)\Gamma(1/-\xi)}(x/(\theta/-\xi))^{1-1}(1-x/(\theta/-\xi))^{\frac{1}{-\xi}-1}(1/(\theta/-\xi))$$

$$= f_{(\theta/-\xi)\cdot Beta\left(1,\frac{1}{-\xi}\right)}(x)$$

$$\Rightarrow \quad \frac{X}{(\theta/-\xi)} \sim Beta\left(1,\frac{1}{-\xi}\right) \sim \frac{Gamma(1,1)}{Gamma(1,1)+Gamma\left(1,\frac{1}{-\xi}\right)}$$

Therefore, depending on the shape parameter, the tail of this distribution can be finite ("no-tailed" $\xi < 0$), medium-tailed ($\xi = 0$), or very heavy-tailed ($\xi > 0$).

What is most relevant to our analysis of tail characteristics is that this distribution is an exponential transformation of the exponential distribution:

$$\frac{X}{\theta} \sim \frac{e^{\xi \cdot Exponential(1)} - 1}{\xi}$$

Under this transformation the light-tailed exponential distribution becomes very heavy-tailed.

Pricing Catastrophe Excess of Loss Reinsurance using Market Curves

David Morel, ACAS

Abstract: What is a simple way to price a catastrophe excess of loss reinsurance program (Cat XL)? By simple we mean pricing a Cat XL with limited information. This paper presents pricing methods that only require the layer pricing of last year's Cat XL program and do not require any catastrophe modelling output.

The first method is to fit a power curve (i.e. a market curve) through the midpoints of the original Cat XL layers and then using that power curve to price the new program. This method has a history of actual use in the reinsurance market.

However, power curves have three key weaknesses and we therefore propose a new method. In this new method we propose a more sophisticated spline curve as the market curve, and unlike the power curve, layers are not represented by their midpoints, but rather by integrating from one endpoint to another. We show how this spline method resolves the three weaknesses of the power curve method.

Note:

An Excel workbook accompanies this paper. There are tabs numbered from #1 to #10. We invite the reader to follow along in the workbook as instructed in the paper so as to increase his or her understanding of the methods. In the workbook, cells that serve as user inputs are highlighted in green. The parameters of market curves (power curves and splines) and the outputs of those market curves are shown in blue.

There are three graphs presented in the workbook that correspond with the three graphs presented in this paper. Should the reader wish to use his or her own Cat XL program in the workbook the axes of the graphs may need to be modified.

1. MOTIVATION OF THE PROBLEM

To motivate the problem let's assume a catastrophe excess of loss reinsurance program (abbreviated in this paper as Cat XL) for a fictional insurance company called Island Insurance. Island Insurance writes property insurance exclusively on a small island with exposure to catastrophic perils such as hurricanes and earthquakes. The total insured value of all of Island Insurance's policies (abbreviated in this paper as TIV) adds up to $2.7 billion USD. Island protects itself from the catastrophic perils with a Cat XL program as follows:

Table 1 - Original Program				
Total Insured Value - TIV		2,700,000,000		
Layer i	Limit - L	Deductible - D	ROL - R	C = L x R Cost - C
Layer 1	5,000,000	5,000,000	20.70%	1,035,000
Layer 2	10,000,000	10,000,000	14.55%	1,455,000
Layer 3	30,000,000	20,000,000	10.20%	3,060,000
Layer 4	50,000,000	50,000,000	6.42%	3,210,000
Layer 5	55,000,000	100,000,000	3.75%	2,062,500
Total Program	**150,000,000**	**5,000,000**	**7.22%**	**10,822,500**

Some comments:

- ROL = Rate on Line = upfront cost of reinsurance layer / Limit of Layer
- Cost = Limit of Layer x ROL
- We ignore reinstatements by assuming that all layers are purchased with the same reinstatement conditions.
- The green cells are user inputs. We recommend that the reader follow along by opening the blank workbook that accompanies this paper, select tab #1 and fill in TIV = 2,700,000,000 in cell D3 and the appropriate Limits, Deductibles and ROLs in columns C, D and E. Note that only the first deductible is necessary in cell D6.

Let us now say that for the following year, Island's TIV went up from $2.7B to $3.0B and also they are restructuring the program into four layers: $7.5m xs $7.5m, $20m xs $15m, $50m xs $35m and $90m xs $85m (for a total program of $167.5m xs $7.5m, thus increasing their total limit from $150m to $167.5m and their retention from $5m to $7.5m).

The question that this paper attempts to answer is straightforward - what do we expect the new market ROLs to be for the new program layers if we don't have any addition information? The only information we have at our disposal is last year's Cat XL program and TIV (all given in Table 1) and this year's proposed Cat XL program and new TIV. We do not have catastrophe modelling information.

Although the TIV is changing year over year, we will otherwise be assuming "flat" renewal conditions:

- Underlying mix of business stays the same
- Geographic footprint stays the same
- Reinsurance market is neither hardening nor softening

What we want is a starting point for the Cat XL renewal.

2. CURRENT SOLUTION: FITTING A POWER CURVE

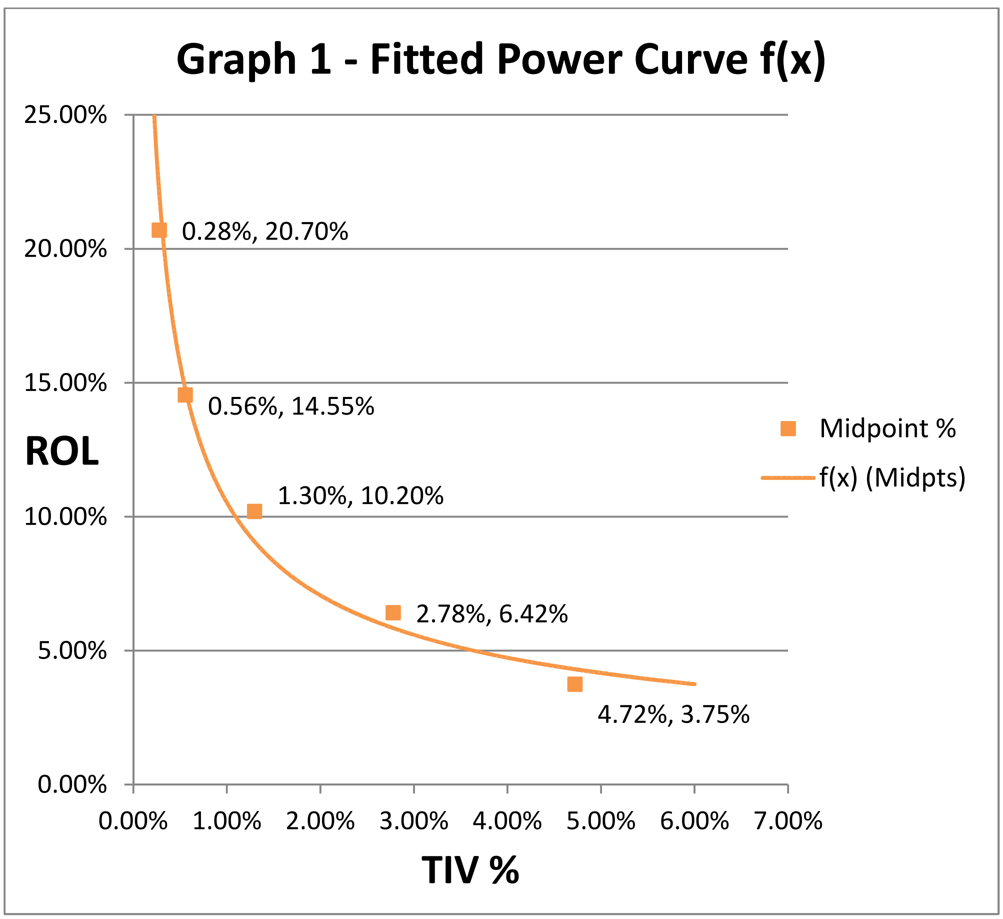
The current solution to the problem posed above is to fit a power curve through the midpoints of the original program layers. This method of fitting a power curve has been known to participants in the London reinsurance market since at least the early 1990s; however, no published document presenting this method has been found by the author of this paper.

In particular, the current solution is as follows: for each layer in the original program, calculate the midpoint of the layer as a % of the original TIV. These are the x values. By midpoint we are referring here to the arithmetic mean or simple average so that if the limit of the layer is L and the deductible of the layer is D, then $x = \frac{AVG(D, D+L)}{TIV}$.

The y values are the ROLs of the layers. Let $f(x) = y = a * x^{-b}$ be a power curve through the points (x, y). We take the logarithm of both sides, and get $\ln(y) = \ln(a) - b * \ln(x)$. ln(y) and ln(x) are thus related linearly, and we calculate a and b to minimize the SSE between the left hand side and the right hand side of the equation.

For Island Insurance, we calculate a = 0.00742 and b = 0.57591. These parameters can be found in blue on tab #2. The actual regression formulas can be found in the hidden columns K and L.

Graphically the power curve looks as follows:



We now use f(x) to price out the new program. In tab #4 we can enter in the new TIV of 3,000,000,000 in cell C3 and the new layering (four new layers) in columns C and D. The new midpoints as a % of the new TIV are calculated in column E, and f(x) is applied to these midpoints to get the new ROLs in column F.

The result is as follows:

Table 2 - New Program Layering: Priced using Power Curve f(x)					
New TIV	3,000,000,000				
			AVG(D, D+L) / TIV	**f(MP)**	**L x ROL_1**
Layer i	**Limit - L**	**Deductible - D**	**Midpoint % - MP**	**ROL_1**	**$Cost_1$**
Layer 1	7,500,000	7,500,000	0.38%	18.51%	1,388,155
Layer 2	20,000,000	15,000,000	0.83%	11.69%	2,337,163
Layer 3	50,000,000	35,000,000	2.00%	7.06%	3,529,088
Layer 4	90,000,000	85,000,000	4.33%	4.52%	4,069,582
Total	**167,500,000**	**7,500,000**		**6.76%**	**11,323,987**

Example Calculation 1 –

Let's calculate the cost of layer #2 of the new program. This layer is \$20m xs \$15m. Given a TIV of \$3B, the midpoint % is x = $\frac{(\$35m+\$15m)/2}{\$3b}$ = 0.833%. Then we calculate the ROL = $f(x) = a * x^{-b} = 0.00742 \times (0.833\%)^{-0.57591}$ = 11.69%. So the cost of the layer = ROL x Limit = 11.69% x \$20m = \$2.34m. (This calculation can be found in cells E7 and F7 on tab #4.)

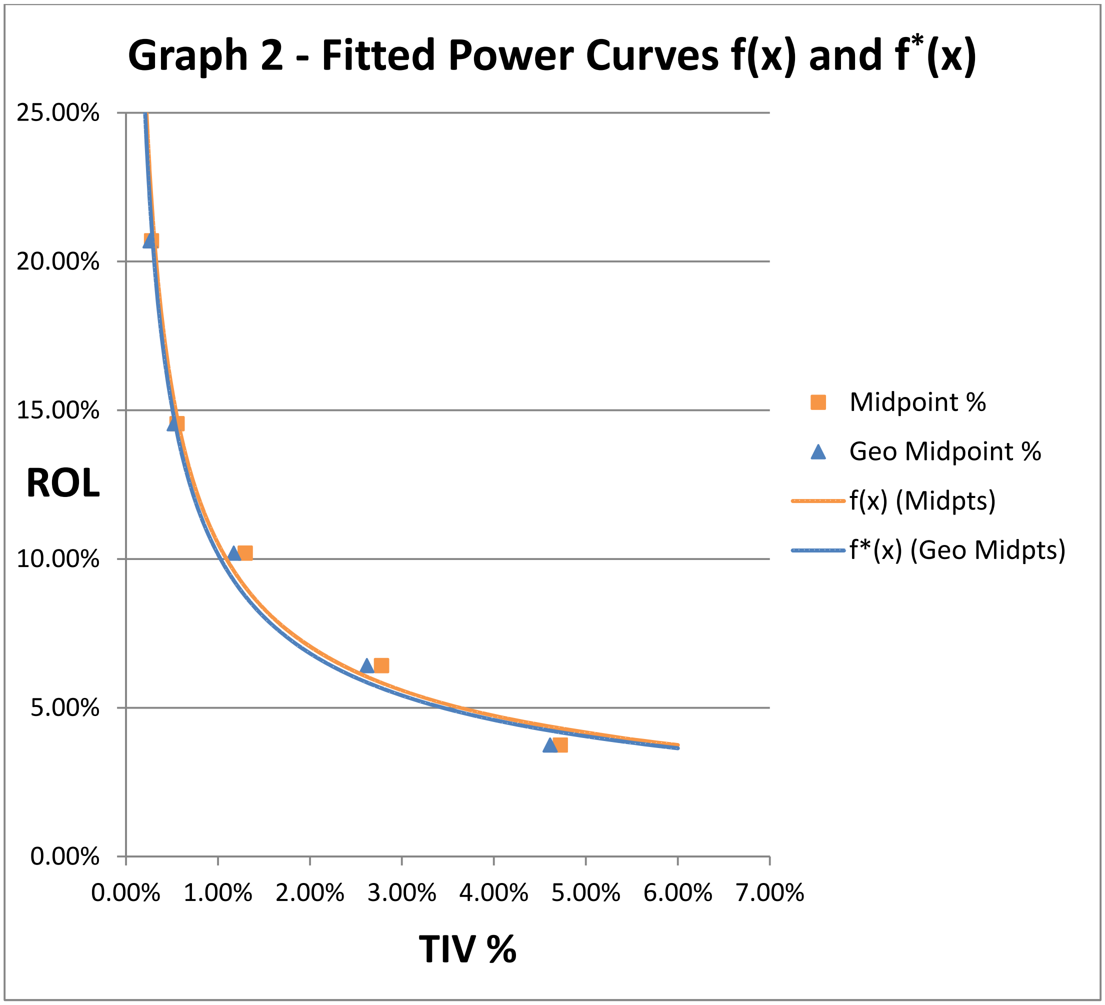
3. POWER CURVE USING GEOMETRIC MIDPOINTS

The solution above we might refer to as the power curve method using midpoints. However, instead of arithmetic midpoints (arithmetic mean) we could also take the geometric mean of each layer to get slightly different results.

Similar to before, for each layer in the original program we calculate the geometric midpoint of the layer as a % of the original TIV. These are the new x values. For a layer limit L and a layer deductible D, we have $x = \frac{\sqrt{D*(D+L)}}{TIV}$. The y values are the same as before – the ROLs of the layers. Let $f^*(x) = y = a * x^{-b}$ be a new power curve through the new points (x, y). We can once again calculate a and b after taking the logarithm of both sides and solving the linear regression.

For Island Insurance, we calculate a = 0.00727 and b = 0.57264. These parameters can be found in blue on tab #5. The actual regression formulas can be found in the hidden columns K and L.

Graphically we can look at both power curves side by side:



We now use $f^*(x)$ to price out the new program. In tab #7 the new TIV of 3,000,000,000 and the new Cat XL program structure are passed over from tab #4 and the calculated ROLs using both f(x) and $f^*(x)$ are shown in columns G and H respectively.

The result is as follows:

Table 3 - New Program Layering: Priced using Power Curves f(x) and $f^*(x)$						
New TIV	3,000,000,000					
			AVG(D, D+L) / TIV	**SQRT[D x (D+L)] / TIV**	**f(MP)**	**f^*(GMP)**
Layer i	**Limit - L**	**Deductible - D**	**Midpoint % - MP**	**Geo Midpoint % - GMP**	**ROL_1**	**ROL_2**
Layer 1	7,500,000	7,500,000	0.38%	0.35%	18.51%	18.42%
Layer 2	20,000,000	15,000,000	0.83%	0.76%	11.69%	11.85%
Layer 3	50,000,000	35,000,000	2.00%	1.82%	7.06%	7.21%
Layer 4	90,000,000	85,000,000	4.33%	4.07%	4.52%	4.55%
Total	**167,500,000**	**7,500,000**			**6.76%**	**6.84%**

Example Calculation 2 –

Let's calculate the cost of layer #2 of the new program using $f^*(x)$. This layer is \$20m xs \$15m. Given a TIV of \$3B, the geometric midpoint % is x = $\frac{\sqrt{\$15m * \$35m}}{\$3b}$ = 0.764%. Then we calculate the ROL = $f^*(x) = a * x^{-b} = 0.00727 \times (0.764\%)^{-0.57264}$ = 11.85%. So the cost of the layer = ROL x Limit = 11.85% x \$20m = \$2.37m. (This calculation can be found in cells F7 and H7 on tab #7.)

What happened when we used geometric midpoints? In *this* case, as we can see in Table 3, we have higher ROLs for Layers 2, 3 and 4 and a lower ROL for Layer 1. Overall the program pricing is higher at 6.84% ROL using geometric midpoints than 6.76% ROL using arithmetic midpoints. We stress that these pricing differences are for Island Insurance only, and the author has not found a general rule as to when geometric midpoints lead to higher prices than arithmetic midpoints and vice versa.

Having now looked at fitting power curves to both types of midpoints, we might now naturally ask, which type of midpoint is better? In the author's practice arithmetic midpoints are used first because they are simpler to understand, and geometric midpoints are used second as a complement to the arithmetic midpoints (if at all).

However, using geometric midpoints may have a theoretical justification. Note that the geometric midpoint of a given layer is always smaller than (to the left of) the arithmetic midpoint. Furthermore, since the power curve is a decreasing function, the "weighted" midpoint of a layer will also be to the left of the arithmetic midpoint.

4. WEAKNESSES OF POWER CURVES

Weakness #1 (Pricing of Original Program) – To see the first weakness of the power curve method let's price out the original Cat XL program for Island Insurance on f(x), which can be seen on tab #2:

Table 4 - Fit a Power Curve through Original ARITHMETIC Midpoints					Fitted ROLs		
Total Insured Value - TIV	2,700,000,000						
			AVG(D, D+L) / TIV		**r = f(MP)**	**c = L x r**	**(r - R) / R**
Layer i	**Limit - L**	**Deductible - D**	**Midpoint % - MP**	**ROL - R**	**ROL - r**	**Cost - c**	**Error %**
Layer 1	5,000,000	5,000,000	0.28%	20.70%	22.00%	1,100,036	6.3%
Layer 2	10,000,000	10,000,000	0.56%	14.55%	14.76%	1,475,949	1.4%
Layer 3	30,000,000	20,000,000	1.30%	10.20%	9.06%	2,718,141	-11.2%
Layer 4	50,000,000	50,000,000	2.78%	6.42%	5.84%	2,920,782	-9.0%
Layer 5	55,000,000	100,000,000	4.72%	3.75%	4.30%	2,366,871	14.8%
Total Program	**150,000,000**	**5,000,000**		**7.22%**	**7.05%**	**10,581,778**	**-2.2%**

Here we have the original TIV of \$2.7B and the original layering, yet when we apply f(x) to the midpoint %'s we get ROLs that are different from the original ROLs. In some cases the error % is high; for the third layer f(x) is underestimating the ROL by 11.2%, for the fifth layer f(x) is overestimating the ROL by 14.8%.

These errors can also be seen by looking at the power curve in Graph 1 – notice that the curve does not go precisely through the points (the actual ROLs), some are above the curve and some are below. Similar error %s can be found for $f^*(x)$ on tab #5.

Naturally, whatever pricing method we choose, we would want the new prices (the starting point) to be the same as the old prices if nothing has changed. The power curve method does not have this important desired property.

Weakness #2 (Non Uniqueness of Layers) – To see the second weakness of the power curve method let's consider the following four distinct layers (Limit L xs Deductible D):

- $1m xs $12m
- $5m xs $10m
- $10m xs $7.5m
- $15m xs $5m

Notice that the midpoint of each of the above layers is $12.5m (Midpoint = AVG(D, D+L) = D + L/2), meaning that under the power curve method (using arithmetic midpoints), each of the above layers would be assigned the same ROL under f(x). Many other layers could be generated.

While we might expect the ROLs for some of these layers to be similar or even the same, there is no reason to believe that *all* of these layers *must* have the same ROLs, as required by the power curve method, so we can consider this a weakness.

Weakness #3 (Unboundedness) – Notice that the power curves are unbounded. $f(x) = a * x^{-b}$ goes to infinity as x goes to 0. This means that if we use a power curve to price layers excess of 0 (i.e. Cat XL layers with no deductible), then the ROL of these layers will get arbitrarily large as the midpoint approaches 0. We will eventually have ROLs (e.g. 1,000%) that do not make sense.

5. PROPOSED SOLUTION: SPLINE CURVE

The power curves above allow the user to find the "market price" of a given Cat XL layer. Thus, in a more general sense, we might refer to these power curves as market curves, and we might also expect to find other, different market curves.

The new market curve proposed in this section is the use of a spline, fitted to the original Cat XL program.

Our first step is to re-envision the way the curves are used to calculate the premium cost of a layer. Instead of getting the midpoint % for the layer and calculating the ROL using f(x) or $f^*(x)$, as we have been doing with the power curve method, let's instead use integration, and integrate from one endpoint of the layer to the other:

Example Calculation 3 –

Let's calculate the cost of the same layer #2 in the new program, the layer \$20m xs \$15m, as we have in example calculations 1 and 2, but this time by integrating f(x) across the layer. First let's define the endpoints of the layer. Let LP_2 = \$15m / \$3b = 0.5% be the left endpoint as a % of TIV, and let RP_2 =\$35m / \$3b = 1.167% be the right endpoint as a % of TIV. Integrating f(x) from LP_2 to RP_2, we have cost = $\int_{LP_2}^{RP_2} a * x^{-b} dx = \frac{a}{-b+1} * \left(x^{-b+1} \Big|_{LP_2}^{RP_2} \right)$.

Plugging in a = 0.00742 and b = 0.57591, we get cost = $\frac{0.00742}{-0.57591+1} \times \left(1.167\%^{-0.57591+1} - 0.5\%^{-0.57591+1}\right)$ = 0.08%.

Since we are integrating across x, and x is expressed as a % of the TIV, this cost is also a % of the TIV. So the cost in dollars would be 0.08% x \$3b = \$2.40m. Finally, the ROL can be worked out as ROL = Cost / Limit = \$2.40m / \$20m = 12.0%.

Notice that the use of integration to calculate the layer costs automatically resolves weakness #2 of the power curve method. That is to say, layers with the same midpoints do not necessarily yield the same ROLs under integration. That is because layers are uniquely defined by their two endpoints, and since integration happens from one endpoint of a layer to the other endpoint, each layer has a unique integration.

The second step of the proposal is to pick a price curve that improves upon f(x). Let's call this new price curve g(x). We would want to pick a g(x) that has the following features:

- Resolves weakness #1 of the power curve. In other words, if we use g(x) to price the original program, we should get the original ROLs.
- The function should be bounded on the top and on the bottom. By bounding the function on the top, as it goes to 0, we resolve weakness #3. By bounding on the bottom, we have a chance to incorporate market knowledge that is external to the Cat XL program itself. For example, we might make the assumption that reinsurers will never price a layer at less than 1% ROL, no matter the underlying exposure. This information is not incorporated into f(x) but we could incorporate it into g(x).

Let n – 1 = the number of layers in the original Cat XL program. (Island Insurance has 5 layers in the original program, so n – 1 = 5). Then let g(x) be a spline with n + 1 segments (so the Island spline will have 7 segments). Let the first segment be linear, the next n – 1 segments be quadratic (representing the original layers) and the last segment be linear. Such a g(x) can be constructed in a way that resolves weakness #1 and is bounded on the top and on the bottom with a maximum ROL and a minimum ROL.

How do we do this? First, let us write down the equations for g(x). We count the n + 1 segments as 0, 1, ..., n, (n = 6 for Island Insurance). Then segments 0 and n are linear and segments 1, 2, ..., n – 1 are quadratic (these correspond to the n – 1 layers in the original Cat XL program). Let's denote the formula for segment i (or layer i) as $g_i(x)$ where $g_i(x)$ is defined on the interval (LP_i, RP_i). Then we have:

- $g_i(x) = a_i + b_i x$ for i = 0 and i = n (first and last segments are linear)
- $g_i(x) = a_i + b_i x + c_i x^2$ for i = 1, 2, ..., n – 1 (middle layers are quadratic)

Where:

- $LP_0 = 0$
- $LP_i = \frac{D_i}{TIV}$ (left endpoint)
- $RP_i = \frac{D_i + L_i}{TIV}$ (right endpoint)
- $RP_i = LP_{i+1}$ (endpoints are connected)
- D_i is the deductible of the i-th layer of the original Cat XL program (i = 1, 2, ... , n – 1)
- L_i is the limit of the i-th layer of the original Cat XL program (i = 1, 2, ... , n – 1)
- RP_n is a point beyond the original Cat XL program (maximum right endpoint)

How do we pick the coefficients a_i, b_i, c_i for g(x)? We want the following conditions to hold:

Condition #1 (Continuity): We want g(x) to be a continuous function; that is, we want the $g_i(x)$ to be connected at the endpoints. Here we have the equations $g_i(RP_i) = g_{i+1}(LP_{i+1})$ for i = 0, 1, ..., n – 1.

Condition #2 (Smoothness): We also want the first derivative $g'(x)$ to be continuous. In other words, we want the function g(x) to be smooth. This is a necessary condition for g(x) to be considered a quadratic spline. Here we have the equations $g_i^{'}(RP_i) = g_{i+1}^{'}(LP_{i+1})$ for i = 0, 1, ..., n – 1.

Condition #3 (Integration): We want for i = 1, 2, ..., n – 1 that $\int_{LP_i}^{RP_i} g_i(x)dx = p_i$ where $p_i = \frac{Cost_i}{TIV}$ and $Cost_i = ROL_i \times L_i$. $Cost_i$, ROL_i and L_i are the known Cost, ROL and Limit of the i-th layer of the original Cat XL program.

This will immediately resolve weakness #1 of the power curve as we are in essence "forcing" g(x) to integrate over the original layers to the original prices.

Condition #4 (Maximum): We want to bound g(x) on the top. We let g(0) = $g_0(0) = a_0 + b_0 \times 0 = a_0 = ROL_{MAX}$. This resolves weakness #3.

Condition #5 (Minimum): We want to bound g(x) on the bottom. We let $g(RP_n) = g_n(RP_n) = a_n + b_n \times RP_n = ROL_{MIN}$

Note that for conditions 4 and 5 the user is required to make a selection for the variables ROL_{MAX}, ROL_{MIN} and RP_n. ROL_{MAX}, the maximum possible ROL, would be the ROL charged for a theoretical layer with no deductible and infinitesimal limit. ROL_{MIN} is the lowest possible ROL, which is reached at some point RP_n which lies beyond the limit of the original Cat XL program. How do we make these selections? This is a highly judgmental step. Here are some ideas:

- We could look at Cat XL programs for companies similar to the one we are pricing (if available) and take into consideration the max and min for those programs.
- We could set RP_n as the point beyond which no coverage would ever actually be purchased.

- For ROL_{MIN} we could consider the values taken on by the power curves at RP_n (i.e. $f(RP_n)$ and $f^*(RP_n)$).
- We could simply look at the curve visually and see what selections make it the "smoothest".

We now provide all of the known variables for Island Insurance in a table, the variables that we will need to set up the equations in Conditions 1 - 5:

Table 5 - Known Variables to Solve for Island Insurance Spline						
TIV	2,700,000,000					
			LP = D / TIV	**RP = (D + L) / TIV**		**(L x R) / TIV**
Layer i	**Limit - L**	**Deductible - D**	**Left Endpt % - LP**	**Right Endpt % - RP**	**ROL - R**	**Cost % - p**
Layer 0	*5,000,000*	*0*	*0.00%*	*0.19%*	*n/a*	*n/a*
Layer 1	5,000,000	5,000,000	0.19%	0.37%	20.70%	0.038%
Layer 2	10,000,000	10,000,000	0.37%	0.74%	14.55%	0.054%
Layer 3	30,000,000	20,000,000	0.74%	1.85%	10.20%	0.113%
Layer 4	50,000,000	50,000,000	1.85%	3.70%	6.42%	0.119%
Layer 5	55,000,000	100,000,000	3.70%	5.74%	3.75%	0.076%
Layer 6	*7,000,000*	*155,000,000*	*5.74%*	*6.00%*	*n/a*	*n/a*
Total	**150,000,000**	**5,000,000**			**7.22%**	**0.401%**

Once again, for Island Insurance there are 5 layers in the original program and n = 6. We invite the reader to inspect this table in the workbook on tab #8.1. For Island Insurance, we make the following selections:

- Let RP_n = 6.00% (cell F15)
- Let ROL_{MAX} = 40.00% (cell G18)
- Let ROL_{MIN} = 3.00% (cell G19)

It may be instructive for the reader to inspect the calculation of the endpoints (columns E and F) as well as the calculation of the p_i (column H) in Excel.

6. SOLVING THE SPLINE CURVE PARAMETERS

Our goal now is to use the variables in Table 5 to set up the equations from Conditions 1-5. We will then use the system of equations to solve for the coefficients a_i, b_i, c_i and thus solve g(x).

First some notes on counting the number of equations:

- We have 3n + 1 equations. The continuity equations (from Condition 1) provide n equations (if there are n + 1 segments then there are n equations between the segments). The smoothness equations (from Condition 2) also provide n equations. The integration equations (from Condition 3) provide n – 1 equations (as there are n – 1 original layers). Finally the boundedness equations (from Conditions 4 and 5) provide 2 equations. Adding them all up we get a grand total of n + n + (n – 1) + 2 = 3n + 1 equations.

- Thus for Island Insurance (n = 6) we have 19 equations: 6 equations for the continuity between the 7 segments, 6 equations for the smoothness between the 7 segments, 5 equations so that g(x) integrate to the original Cat XL prices on each original layer, and 2 equations for the boundedness conditions.
- All of the equations are linear. Once we plug in the knowns ($LP_i, RP_i, p_i, ROL_{MAX}, ROL_{MIN}$) then the equations all reduce to linear equations with a_i, b_i, c_i as the unknowns.
- Counting up the number of unknown variables in the 3n + 1 equations we also have 3n + 1 unknowns. $g_0(x)$ and $g_n(x)$ each have 2 unknown coefficients (a_i, b_i), and each of the n – 1 $g_i(x)$ has 3 unknown coefficients (a_i, b_i, c_i), for a grand total of 2 + 2 + 3 x (n – 1) = 3n + 1 unknowns.

Thus we have a system of 3n + 1 linear equations with 3n + 1 unknown variables (the coefficients), allowing us to use matrix algebra to solve for those coefficients.

What exactly do the 19 linear equations look like for Island Insurance? We present them in the following table:

Table 6 - 19 Linear Equations for Island Insurance

Eqn #	Condition / Equation Description		General Form	Expanded Form
1	1	Continuity b/w layers 0 & 1	$g_0(RP_0) = g_1(LP_1)$	$a_0 + b_0*RP_0 = a_1 + b_1*LP_1 + c_1*LP_1^2$
2	1	Continuity b/w layers 1 & 2	$g_1(RP_1) = g_2(LP_2)$	$a_1 + b_1*RP_1 + c_1*RP_1^2 = a_2 + b_2*LP_2 + c_2*LP_2^2$
3	1	Continuity b/w layers 2 & 3	$g_2(RP_2) = g_3(LP_3)$	$a_2 + b_2*RP_2 + c_2*RP_2^2 = a_3 + b_3*LP_3 + c_3*LP_3^2$
4	1	Continuity b/w layers 3 & 4	$g_3(RP_3) = g_4(LP_4)$	$a_3 + b_3*RP_3 + c_3*RP_3^2 = a_4 + b_4*LP_4 + c_4*LP_4^2$
5	1	Continuity b/w layers 4 & 5	$g_4(RP_4) = g_5(LP_5)$	$a_4 + b_4*RP_4 + c_4*RP_4^2 = a_5 + b_5*LP_5 + c_5*LP_5^2$
6	1	Continuity b/w layers 5 & 6	$g_5(RP_5) = g_6(LP_6)$	$a_5 + b_5*RP_5 + c_5*RP_5^2 = a_6 + b_6*LP_6$
7	2	Smoothness b/w layers 0 & 1	$g_0'(RP_0) = g_1'(LP_1)$	$b_0 = b_1 + 2*c_1*LP_1$
8	2	Smoothness b/w layers 1 & 2	$g_1'(RP_1) = g_2'(LP_2)$	$b_1 + 2*c_1*RP_1 = b_2 + 2*c_2*LP_2$
9	2	Smoothness b/w layers 2 & 3	$g_2'(RP_2) = g_3'(LP_3)$	$b_2 + 2*c_2*RP_2 = b_3 + 2*c_3*LP_3$
10	2	Smoothness b/w layers 3 & 4	$g_3'(RP_3) = g_4'(LP_4)$	$b_3 + 2*c_3*RP_3 = b_4 + 2*c_4*LP_4$
11	2	Smoothness b/w layers 4 & 5	$g_4'(RP_4) = g_5'(LP_5)$	$b_4 + 2*c_4*RP_4 = b_5 + 2*c_5*LP_5$
12	2	Smoothness b/w layers 5 & 6	$g_5'(RP_5) = g_6'(LP_6)$	$b_5 + 2*c_5*RP_5 = b_6$
13	3	Area - Layer 1	$\int g_1(x) = p_1$	$a_1*(RP_1 - LP_1) + ½*b_1*(RP_1^2 - LP_1^2) + ⅓*c_1*(RP_1^3 - LP_1^3) = p_1$
14	3	Area - Layer 2	$\int g_2(x) = p_2$	$a_2*(RP_2 - LP_2) + ½*b_2*(RP_2^2 - LP_2^2) + ⅓*c_2*(RP_2^3 - LP_2^3) = p_2$
15	3	Area - Layer 3	$\int g_3(x) = p_3$	$a_3*(RP_3 - LP_3) + ½*b_3*(RP_3^2 - LP_3^2) + ⅓*c_3*(RP_3^3 - LP_3^3) = p_3$
16	3	Area - Layer 4	$\int g_4(x) = p_4$	$a_4*(RP_4 - LP_4) + ½*b_4*(RP_4^2 - LP_4^2) + ⅓*c_4*(RP_4^3 - LP_4^3) = p_4$
17	3	Area - Layer 5	$\int g_5(x) = p_5$	$a_5*(RP_5 - LP_5) + ½*b_5*(RP_5^2 - LP_5^2) + ⅓*c_5*(RP_5^3 - LP_5^3) = p_5$
18	4	Maximum ROL	$g_0(LP_0) = ROL_{MAX}$	$a_0 + b_0*LP_0 = a_0 = ROL_{MAX}$
19	5	Minimum ROL	$g_6(RP_6) = ROL_{MIN}$	$a_6 + b_6*RP_6 = ROL_{MIN}$

These 19 equations have 19 unknowns: $a_0, b_0, a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3, a_4, b_4, c_4, a_5, b_5, c_5, a_6, b_6$.

These equations can also be found in tab #8.2, although given that the workbook is designed to handle up to 8 original layers, there are some dummy equations listed there too (i.e. $a_7 = b_7 = c_7 = ... = a_9 = b_9 = 0$).

We then convert the 19 equations in Table 6 to the following form:
$k_{j1}a_0 + k_{j2}b_0 + k_{j3}a_1 + k_{j4}b_1 + k_{j5}c_1 + ... + k_{j18}a_6 + k_{j19}b_6 = s_j$.
That is to say, we convert each equation into a linear combination of the unknowns on the left hand side and a solution constant on the right hand side. k_{ji} is a known factor for the j-th equation and the i-th unknown variable. s_j is the solution constant to the j-th equation.

For example, let's take equation 2 which is $a_1 + b_1 RP_1 + c_1 RP_1^2 = a_2 + b_2 LP_2 + c_2 LP_2^2$. From Table 5, we have that $RP_1 = LP_2 = 0.37\%$. Then equation 2 in the prescribed format is as follows:
$a_1 + 0.37\% \times b_1 + 0.001369\% \times c_1 - a_2 - 0.37\% \times b_2 - 0.001369\% \times c_2 = 0$. Note that all the other unknowns have a factor of 0 and are not shown here.

Converting the 19 equations to matrix form, we take the k_{ji} of the left hand side and let $\mathbf{A} = \begin{pmatrix} k_{1,1} & \cdots & k_{1,19} \\ \vdots & \ddots & \vdots \\ k_{19,1} & \cdots & k_{19,19} \end{pmatrix}$.

We also form the unknown equation vector $\mathbf{X} = \begin{pmatrix} a_0 \\ \vdots \\ b_6 \end{pmatrix}$ and the solution vector $\mathbf{B} = \begin{pmatrix} s_1 \\ \vdots \\ s_{19} \end{pmatrix}$.

Thus we have the equation **A * X = B**. Finally, we use matrix algebra to find the inverse of **A**, $\mathbf{A}^{-1}$. Then
$\mathbf{X} = \mathbf{A}^{-1} * \mathbf{B}$ and we have solved for all the unknown coefficients simultaneously, thus solving g(x).
The full matrices **A**, $\mathbf{A}^{-1}$ and the vectors **B** and **X** can all be found on tab #8.2.

The solved coefficients of g(x) – the spline curve for Island Insurance – are summarized in the following table:

Table 7 - Island Insurance Layer Summary and Solved Spline Coefficients

TIV	2,700,000,000								
			LP = D / TIV	**RP = (D + L) / TIV**		**(L x R) / TIV**			
Layer i	**Limit - L**	**Deductible - D**	**Left Endpt % - LP**	**Right Endpt % - RP**	**ROL - R**	**Cost % - p**	**a**	**b**	**c**
Layer 0	*5,000,000*	*0*	*0.00%*	*0.19%*	*n/a*	*n/a*	0.40	-75.83	
Layer 1	5,000,000	5,000,000	0.19%	0.37%	20.70%	0.038%	0.45	-132.95	15422.37
Layer 2	10,000,000	10,000,000	0.37%	0.74%	14.55%	0.054%	0.27	-31.82	1769.90
Layer 3	30,000,000	20,000,000	0.74%	1.85%	10.20%	0.113%	0.18	-7.77	146.95
Layer 4	50,000,000	50,000,000	1.85%	3.70%	6.42%	0.119%	0.13	-3.23	24.29
Layer 5	55,000,000	100,000,000	3.70%	5.74%	3.75%	0.076%	0.14	-3.52	28.22
Layer 6	*7,000,000*	*155,000,000*	*5.74%*	*6.00%*	*n/a*	*n/a*	*0.05*	*-0.28*	
Total	**150,000,000**	**5,000,000**			**7.22%**	**0.401%**			

These solved coefficients can be found on tab #8.1. Let's now check to see if $g(x)$ is working the way we want it to work by calculating the values of $g(x)$ and $g'(x)$ on the segment endpoints, and integrating $g(x)$ across the segments:

Table 8 - Island Insurance Layer Summary and Verification of Properties of Spline Curve

TIV	2,700,000,000		LP =	RP =		(L x R) /					
		Deductible	D / TIV	(D + L) / TIV	ROL	TIV					
Layer i	**Limit - L**	**D**	**Left Endpt**	**Right Endpt**	**R**	**Cost % - p**	**g(LP)**	**g(RP)**	**g'(LP)**	**g'(RP)**	$\int_{LP}^{RP} g(x)$
Layer 0	*5,000,000*	*0*	*0.00%*	*0.19%*	*n/a*	*n/a*	40.00%	25.96%	-75.83	-75.83	n/a
Layer 1	5,000,000	5,000,000	0.19%	0.37%	20.70%	0.038%	25.96%	17.20%	-75.83	-18.71	0.038%
Layer 2	10,000,000	10,000,000	0.37%	0.74%	14.55%	0.054%	17.20%	12.70%	-18.71	-5.60	0.054%
Layer 3	30,000,000	20,000,000	0.74%	1.85%	10.20%	0.113%	12.70%	8.30%	-5.60	-2.33	0.113%
Layer 4	50,000,000	50,000,000	1.85%	3.70%	6.42%	0.119%	8.30%	4.82%	-2.33	-1.43	0.119%
Layer 5	55,000,000	100,000,000	3.70%	5.74%	3.75%	0.076%	4.82%	3.07%	-1.43	-0.28	0.076%
Layer 6	*7,000,000*	*155,000,000*	*5.74%*	*6.00%*	*n/a*	*n/a*	3.07%	3.00%	-0.28	-0.28	n/a
Total	**150,000,000**	**5,000,000**			**7.22%**	**0.401%**					

Checking Condition #1 (Continuity): Notice that the value that $g(x)$ takes at the right of segment i is equal to the value that $g(x)$ takes at the left of segment i + 1. For example, $g_0(RP_0) = g_1(LP_1)$ = 25.96%. This implies that $g(x)$ is continuous.

Checking Condition #2 (Smoothness): Similarly we notice that the value that $g'(x)$ takes at the right of segment i is equal to the value that $g'(x)$ takes at the left of segment i + 1. For example, $g'_0(RP_0) = g'_1(LP_1)$ = -75.83. This implies that $g'(x)$ is continuous (i.e. that $g(x)$ is smooth). Note that the derivative is negative throughout, which means that $g(x)$ is decreasing throughout. Also note that the derivative while negative is also increasing, which means that $g(x)$ is concave up (the power curves f(x) and $f^*(x)$ are also concave up).

Checking Condition #3 (Integration): Notice that the integral of each $g_i(x)$ on its defined interval LP_i to RP_i is equal to p_i (i.e. the last column in Table 8, $\int_{LP_i}^{RP_i} g(x)dx$, is equal to the 7th column in Table 8, p_i). Thus if we use g(x) to price the original program with no change in TIV, we get the same ROLs as the original program.

Checking Conditions #4 and #5 (Maximum and Minimum): Notice that $g_0(LP_0)$ = g(0%) = 40.00% = ROL_{MAX} and $g_6(RP_6)$ = g(6.00%) = 3.00% = ROL_{MIN}.

These verifications can also be found on tab #8.1.

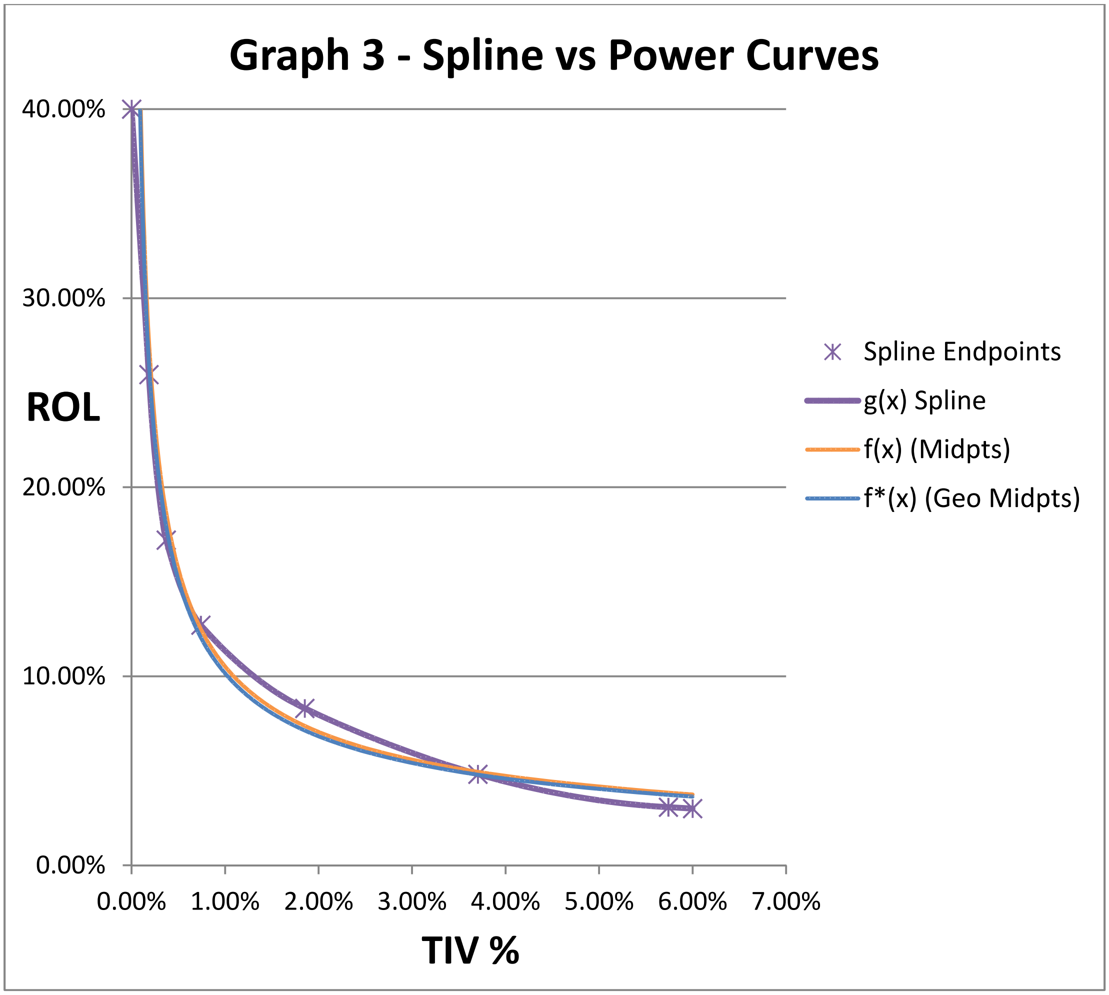
7. COMPARISON OF THE METHODS

We have now solved for f(x), $f^*(x)$ and g(x). Let's use all three of them to price out the new Cat XL program for Island Insurance with the new TIV:

Table 9 - Pricing of New Program using all the Methods

New TIV 3,000,000,000 — New Program Layering			Endpoints and Midpoints				Power Curve f(x)		Power Curve $f^*(x)$		Spline Method	
Layer i	Limit - L	Deductible D	LP	GMP	MP	RP	f(MP) ROL$_1$	L x ROL$_1$ Cost$_1$	f^*(GMP) ROL$_2$	L x ROL$_2$ Cost$_2$	Cost$_3$ / L ROL$_3$	$\int_{LP}^{RP} g(x) * TIV$ Cost$_3$
Layer 1	7,500,000	7,500,000	0.25%	0.35%	0.38%	0.50%	18.51%	1,388,155	18.42%	1,381,650	17.53%	1,314,627
Layer 2	20,000,000	15,000,000	0.50%	0.76%	0.83%	1.17%	11.69%	2,337,163	11.85%	2,370,376	12.37%	2,473,283
Layer 3	50,000,000	35,000,000	1.17%	1.82%	2.00%	2.83%	7.06%	3,529,088	7.21%	3,606,327	8.10%	4,047,793
Layer 4	90,000,000	85,000,000	2.83%	4.07%	4.33%	5.83%	4.52%	4,069,582	4.55%	4,094,577	4.24%	3,813,139
Total	**167,500,000**	**7,500,000**					**6.76%**	**11,323,987**	**6.84%**	**11,452,929**	**6.95%**	**11,648,842**

This table can be found on tab #10. We see that using geometric midpoints, i.e. $f^*(x)$, results in ROLs that are generally higher than using arithmetic midpoints, i.e. f(x). We also see that, for this example, using the spline method results in the highest overall ROL (6.95%), with significant variation from layer to layer. For example, the cost of layer 3 is 14.7% higher under g(x) than under f(x) (8.10% / 7.06%), yet the cost of layer 4 is 6.2% lower (4.24% / 4.52%). This variation from layer to layer is due to parameterized flexibility of the spline curve, which becomes more apparent when we observe it visually:



The spline endpoints marked by asterisks indicate the left and right endpoints of the original Cat XL program, in addition to LP_0 and RP_6. As can be seen, the spline is strictly decreasing throughout and is higher than the power curves on original layers 3 and 4 and lower on original layer 5. We can also clearly see that g(x) is bounded at the top, that g(0) = 40.00%.

We encourage the reader to try pricing out a new Cat XL program in the Excel workbook by entering in the new Cat XL structure in green in tab #4 and then examining the output pricing in tab #10. Here are some structures to test:

- Enter in the original TIV = $2,700,000,000 and the original structure: $5m xs $5m, $10m xs $10m, $30m xs $20m, $50m xs $50m and $55m xs $100m. Compare tab #1 with tab #10. Notice that the layer ROLs are preserved under the spline curve g(x) but are not preserved under the power curves. The correct overall ROL for the program is 7.22%.
- Now go back to tab #4, keep the TIV as it is, and enter the following three layers: $20m xs $5m, $30m xs $25m and $100m xs $55m. This gives an overall program of $150m xs $5m, the same as before. Notice in tab #10 that the overall ROL under the spline method stays the same at 7.22%. This is a property of using the spline method – while individual layer prices will be different, the overall price will be the same because integration is additive: $\int_{LP_1}^{RP_1} g(x)dx + \int_{LP_2}^{RP_2} g(x)dx = \int_{LP_1}^{RP_2} g(x)dx$.

8. ADVANTAGES OF POWER CURVES

While we have presented the spline method as superior to power curves in that it resolves the three key weaknesses in section 4, there are ways in which power curves are superior to splines:

- Power curves are easier to set up and explain than splines.
- Splines require ROL_{MAX} and ROL_{MIN} to be specified judgmentally in the model; the power curves require no such selections.
- The power curve function is $f(x) = a * x^{-b}$ and its derivative is $f'(x) = -a * b * x^{-b-1}$. Since the derivative is negative for all x, the power curve is strictly decreasing. But for spline curves there is nothing in the definition that enforces this property. Upon visual inspection we may occasionally find a spline curve with a region that is not strictly decreasing. In those cases we might be able to "fix" the curve by selecting a different ROL_{MAX} and ROL_{MIN}.
- Finally, an application: let's say we have the Cat XL programs of several similar insurance companies (e.g. competitors of Island Insurance who also write property policies exclusively on the island). We can then fit a power curve through the midpoints of ALL the layers of ALL the Cat XL programs, thus creating a consolidated market curve for the entire island, not just Island Insurance. It is not obvious how to create such a consolidated market curve using a spline.

9. SUMMARY

In this paper we have presented the concept of pricing a catastrophe excess of loss program (Cat XL) using a market curve. Pricing with such a market curve is simple in that it only requires the total insured value (TIV) of the new program to be priced, and a benchmark program (such as last year's Cat XL program), and does not require the use of catastrophe modelling output.
We then presented the simplest market curve, which is the power curve. The power curve fits a function of the form $f(x) = a * x^{-b}$ to the midpoints (arithmetic or geometric) of the benchmark program, and then this curve is used to price out the new program. We showed, however, that the power curve has three key weaknesses.

We then proposed a new market curve, a spline function, and the use of integration instead of taking the midpoints of layers, which resolves the three key weaknesses of the power curve. We showed how to solve for the spline, which involves solving a system of linear equations.

We provided an Excel workbook that allows the reader to test all the methods.

10. FURTHER RESEARCH

The power curve function takes the form $f(x) = a * x^{-b}$, however other variations could be investigated:

- **Power Curve with a constant:** $f(x) = a * x^{-b} + c$ (Notice that ROL_{MIN} = c)
- **Exponential Decay:** $f(x) = a * b^{-x}$ (Notice that ROL_{MAX} = a)
- **Exponential Decay with a constant:** $f(x) = a * b^{-x} + c$ (Notice that ROL_{MAX} = a + c and ROL_{MIN} = c)

In addition, we could investigate some of the simplifying assumptions that we made in the paper:

- The issue of reinstatements could be studied. What happens if different layers have different reinstatement conditions?
- What happens if we assume a known reinsurance market cycle as opposed to "flat", unchanging rates?

Finally, we could try to develop a formula to relate the market curve and market pricing to the underlying catastrophe exposure.

11. ACKNOWLEDGMENTS

The author would like to thank Neil Franklin for teaching him the original power curve method, Kirk Conrad for his guidance and encouragement, and Neil Bodoff for reading the first draft and providing feedback.

12. REFERENCES

As stated earlier in the paper, the author is not aware of any published document that presents the power curve method, despite the fact that it has been used in the reinsurance market since the early 1990s. However, we present two references for the general edification of the reader. The first is the Clark paper "Basics of Reinsurance Pricing" and the second is the "Loss Models" textbook, which provides a background on spline curves (more advanced splines).

[1] D. Clark, "Basics of Reinsurance Pricing," CAS Study Note, 1996.

[2] Klugman, Stuart A., Harry H. Panjer and G.E. Willmot, G.E. *Loss Models: from Data to Decisions*. Wiley Series in Probability and Statistics, 1998.

Reinsurance Arrangements Minimizing the Total Required Capital

Yingjie Zhang, PhD, FCAS
The Cincinnati Insurance Companies
PO Box 145496, Cincinnati, OH 45250
E-mail: yingjie_zhang@cinfin.com

Abstract

Reinsurance reduces the required capital of the primary insurer but increases that of the reinsurer. Capital is costly. All capital costs, including that of the reinsurer, are ultimately borne by primary policyholders. Reducing the total capital of insurers and reinsurers lowers the total capital cost and the total primary policy premium. A reinsurance arrangement is considered optimal if it minimizes the total required capital. This optimal reinsurance is shown to be an attracting equilibrium under price competition. Evidence suggests that there is an inverse relationship between the total required capital and the correlation between the losses held by different insurers. Examples are constructed to support this observation.

Keywords

Required capital, capital cost, optimal reinsurance, subadditive risk measure, correlation between losses

Acknowledgements

I thank Neb Bojer, Evan Glisson, Art Zaremba and Ka Chun Yeung for reviewing the paper. They have helped me correct errors and improve presentation.

1 Introduction

A new type of optimal reinsurance is introduced in this paper. Reinsurance serves many purposes, one of which is to reduce the required capital by lessening the volatility of losses. From the shareholder point of view, capital is costly because of income taxes and agency costs. Shareholders pay income taxes two times on their capital investment, first at the corporate level and then at the personal level when they sell the stock. They would not owe the corporate tax if they invested directly in the securities market. Agency costs exist because of the separation of ownership and control. They include monitoring and bonding expenditures and other losses in profits due to a misalignment of managers' decisions and shareholders' welfare. Taxes and agency costs, altogether called capital costs, generally have an increasing relationship with the amount of capital (Jensen and Meckling 1976, Perold 2005, Chandra and Sherris 2006, Zhang 2008). Thus carrying less capital is desirable.

Reinsurance transfers losses from a ceding company to a reinsurer. Such losses are often highly volatile. So this transfer of losses increases the capital requirement of a reinsurer while reducing that of a ceding company. Consequently, capital costs of the reinsurer increase and those of the ceding company decrease. The total capital cost, the sum of that of both companies, may go either way. Capital costs are funded by premium. Primary policy premiums include charges to cover primary insurers' capital costs; reinsurance premiums include charges to cover reinsurers' capital costs. But reinsurance premiums are funded through premiums of primary policies. Therefore, the total capital costs of primary insurers and reinsurers are ultimately borne by primary policyholders. If a treaty reduces a ceding company's capital costs more than it increases the reinsurer's, the total capital cost is reduced, which benefits primary policyholders. A treaty, or a set of treaties, is optimal, if it minimizes the total capital cost. Such optimal reinsurance arrangements are the subject of this paper.

Numerous authors have written about optimal reinsurance and have proposed various optimality criteria. My approach is noticeably different. Usually an optimal reinsurance is defined from the ceding company's point of view. The ceding insurer seeks a treaty to maximize its risk-adjusted return (Lampaert and Walhin 2005, Fu and Khury 2010), to minimize the variance of its net loss (Kaluszka 2001, Lampaert and Walhin 2005), or to minimize the tail risk of the net loss (Gajek and Zagrodny 2004, Cai and Tan 2007), under the constraint of a given premium principle that links the ceded premium to the ceded loss. This line of research is valuable. However, it does not pay enough attention to the profit target of the reinsurer. Although the proposed premium principles usually include risk margins reflecting the volatility of the ceded loss, they generally ignore the fact that the reinsurer needs to put up more capital thus incurring greater capital costs. My approach places the ceding insurer and the reinsurer on an equal footing and addresses the capital costs of both directly. A reinsurance arrangement that minimizes the total capital is the best deal for the combined welfare of primary insurers, reinsurers and policyholders.

Under reasonable assumptions, minimization of the total capital cost is equivalent to minimization of the total amount of capital carried by all companies. This latter problem may be directly solved by simulating insurers' and reinsurers' losses. A remarkable fact, however, is that this type of optimal reinsurance need not be solved by any one party. (In fact, neither the ceding insurer nor the reinsurer can obtain the full knowledge of the joint probability distribution of losses of both parties.) Market forces automatically push the insurer and the reinsurer to select treaties with less total capital costs. In other words, an optimal reinsurance arrangement is an attracting equilibrium.

The capital requirement will be set by a risk measure. In this paper, I assume that the risk measure is coherent, as defined in Artzner et al. (1999). For such a risk measure, there is an absolute lower bound for the total capitals. Regardless of reinsurance arrangements, the total capital must be greater than this lower bound. It can be shown that if the losses of the insurers have a certain correlation called comonotonicity (defined in Section 5), then the total capital attains the lower bound. This observation leads to a discussion on the relationship between optimal reinsurance and correlated losses. Evidence suggests that an optimal treaty is one that makes the losses of insurers and reinsurers as correlated as possible. (Such correlation needs only occur at the tail.)

The main part of the paper is organized as follows. In Section 2, I prove that minimization of the total primary insurance premium leads to minimization of the total capital. I then show in Section 3 that price competition tends to produce this type of optimal reinsurance. Coherent risk measures are discussed in Section 4. In Section 5, I point out that a lower bound exists for the total required capital, and in some cases an inverse relationship exists between the sum of capitals and the correlation between losses. Section 6 contains a general formulation of the optimal reinsurance problem. Examples are given in Section 7 to further examine the link between the sum of capitals and correlation. Section 8 concludes the paper.

2 Why Minimize the Total Required Capital?

In this section I will rigorously prove that, if a reinsurance arrangement minimizes the total capital cost, then it minimizes the aggregate premium of primary policyholders. I will also point out the exact conditions under which minimization of the total capital cost is equivalent to minimization of the total amount of capital.

Policyholders purchase insurance to protect themselves against unexpected losses. At the same time, they also provide funds to cover all operating costs of the insurance company, including underwriting and claim expenses, income taxes, agency costs and reinsurance costs. The reinsurance costs, in turn, cover the reinsurer's expenses, taxes and agency costs, and *its* reinsurance costs (costs of retrocession). Ultimately, it is the primary insurance policyholders that bear the operating costs of primary insurers and reinsurers. For the insurance/reinsurance market as a whole, reinsurance treaties rearrange these costs among all insurers and reinsurers. Some reinsurance arrangements result in lower total costs than others. A reinsurance arrangement is optimal if the total cost is minimized, in which case the primary policyholders pay the lowest aggregate premium.

This paper focuses on minimizing the total capital cost, consisting of income taxes and agency costs.[1] To cleanly study the capital cost, I assume that the aggregate underwriting and claim expenses remain constant under various reinsurance arrangements. Therefore, these expenses can be excluded from consideration. The gross insurance premium of a policy can be decomposed into the following components

$$p = \text{PV(Loss)} + \text{PV(Tax)} + \text{PV(Agency Cost)} + \text{Reinsurance Premium.} \quad (2.1)$$

The p in (2.1) represents the fair premium, which is the exact amount to fund all insurer's costs related to the policy. Equation (2.1) is a version of the net present value principle. Slightly different formulas for the fair premium have appeared in the literature (Myers and Cohn 1987, Taylor 1994, Vaughn 1998). Each term on the right-hand side of (2.1) provides the exact amount to cover that specific type of cost. The PV's represent risk-adjusted present values. The loss in the first term is the net loss. It is assumed here that the present value of insured loss satisfies the following two basic requirements of the fair value accounting: (1) The value PV(Loss) is independent of the carrier of the insurance policy.[2] (2) The function PV( ) is additive. The two conditions together eliminate the possibility

[1] Agency costs include any cost associated with the issue of "separation of ownership and control", as discussed in Jensen and Meckling (1976), Perold (2005), like monitoring and bonding expenditures and other losses in profits due to misalignment of managers' decisions and shareholders' welfare.

[2] The risk-adjusted PV can be viewed as the risk-free discounted expected cash flow plus a risk margin, where the risk margin reflects the market, or the systematic risk of the cash flow. It is sometimes argued that the fair value of losses should be affected by its carrier's default risk. In this paper, I only consider insurance firms that hold the required level of capital and whose risk of default is negligible.

of arbitrage. In particular, they imply that PV(Gross Loss) = PV(Net Loss) + PV(Ceded Loss).

I now examine the relationship between the gross fair premium and the total amount of capital held by insurers and reinsurers. Consider a one-year model containing only one loss to be shared between a primary insurer and a reinsurer. Let p be the gross premium charged by the primary insurer at the beginning of the year and L the random gross loss paid at the end of the year. The primary insurer collects the premium p then cedes an amount p_c to the reinsurer, retaining $p_n = p - p_c$. Similarly for losses, $L_n = L - L_c$, where L_c is the ceded loss and L_n the net loss.

The total income tax is the sum of two charges, one on the income generated by premiums, which equals the underwriting profit plus the investment income on premiums, and the other on the investment income generated by capital. To write premium formulas in a concise way, I use the following notations

e_{Pr}	:	capital carried by the primary insurer
e_{Re}	:	capital carried by the reinsurer
t_{Pr}	:	average tax rate for the primary insurer
t_{Re}	:	average tax rate for the reinsurer

The present value of tax for the primary insurer is of the form $t_{Pr}(p_n - \text{PV}(L_n)) + u_{Pr}e_{Pr}$, and that for the reinsurer is $t_{Re}(p_c - \text{PV}(L_c)) + u_{Re}e_{Re}$, where the u's are constants: if r_f represents the risk-free rate, then $u_{Pr} = t_{Pr} \cdot r_f/(1+r_f)$ and $u_{Re} = t_{Re} \cdot r_f/(1 + r_f)$. (A derivation of the multiplier $r_f/(1 + r_f)$ can be found in Cummins 1990). Agency costs generally increase with the amount of capital.[3] For simplicity, I assume there is a linear relationship: for some constants s_{Pr} and s_{Re}, the present value of agency cost is $s_{Pr}e_{Pr}$ for the primary company and $s_{Re}e_{Re}$ for the reinsurer.

Following (2.1), for the primary insurer, we have

$$p = \text{PV}(L_n) + t_{Pr}(p_n - \text{PV}(L_n)) + u_{Pr}e_{Pr} + s_{Pr}e_{Pr} + p_c \tag{2.2}$$

and, for the reinsurer (if there is no retrocession),

$$p_c = \text{PV}(L_c) + t_{Re}(p_c - \text{PV}(L_c)) + u_{Re}e_{Re} + s_{Re}e_{Re}. \tag{2.3}$$

[3] An important type of capital cost is the cost of financial distress, which increases as capital becomes more insufficient. But firms considered in this paper satisfy a given capital requirement. So the cost of financial distress is ignored.

An equation for the fair gross premium p can be obtained by substituting (2.3) into (2.2). p is the sum of the following four terms.

1. The present value of loss: $\mathrm{PV}(L_n) + \mathrm{PV}(L_c) = \mathrm{PV}(L)$, which does not vary with reinsurance.

2. The tax on the incomes generated by premium: $t_{Pr}(p_n - \mathrm{PV}(L_n)) + t_{Re}(p_c - \mathrm{PV}(L_c))$. On the condition that the tax rates are equal, $t_{Pr} = t_{Re} = t$, this term is $t(p - \mathrm{PV}(L))$, which decreases as p decreases.

3. The tax on the incomes generated by capital: $u_{Pr}e_{Pr} + u_{Re}e_{Re}$. If the applicable tax rates are the same, then $u_{Pr} = u_{Re} = u$, and the term equals $u(e_{Pr} + e_{Re})$, which decreases if a reinsurance contract lowers the sum of capitals, $e_{Pr} + e_{Re}$.

4. The agency cost: $s_{Pr}e_{Pr} + s_{Re}e_{Re}$. If the cost factors are equal, $s_{Pr} = s_{Re} = s$, then the term equals $s(e_{Pr} + e_{Re})$, again a direct function of the total capital $e_{Pr} + e_{Re}$.

To sum up, as reinsurance varies, the loss component $\mathrm{PV}(L)$ remains constant, while the fair premium p varies because taxes and agency costs vary. p is lower if the present values of taxes and agency costs are lower. Under the above assumptions, this is equivalent to a less amount of total capital, $e_{Pr} + e_{Re}$. The optimal reinsurance is then defined as the one that minimizes $e_{Pr} + e_{Re}$. An optimal treaty creates the least gross premium, so is best for the policyholder.

This definition can be generalized to an insurance market with many primary insurers and reinsurers, and many primary policyholders. Assume each primary insurer covers a given set of policyholders. There are a great number of ways in which each insurer buys reinsurance and each reinsurer enters retrocession agreements. A set of reinsurance/retrocession arrangements is called optimal if it minimizes the total capital cost of the insurers and reinsurers. With the condition that all companies have identical tax rates and agency cost factors, this criterion is equivalent to minimizing the total amount of capital.[4]

[4] It has been pointed out to me that reinsurers usually have a different tax rate than primary companies. If tax rates or agency cost factors are not all equal, or the costs are not all linear to the capital, then the optimal treaty is one that minimizes some increasing function of the capitals.

3 Market Competition Produces Lower Total Capital

Minimization of the total capital cost is a new optimality criterion. Criteria in the existing literature are very different; see Kaluszka (2001), Gajek and Zagrodny (2004), Lampaert and Walhin (2005), Cai and Tan (2007) and Fu and Khury (2010) for a sample of recent papers. In these papers, reinsurance is considered optimal if it minimizes the risk of the net loss under a given constraint on the reinsurance cost (or a constraint on the ceded premium). This line of research is valuable for reinsurance purchase decisions but is incomplete. A major concern of reinsurance has been missing. The reinsurer needs additional capital to accommodate the increased risk from assumed losses, which increases its capital cost. This extra cost is transferred to the ceding company through reinsurance pricing. To the ceding company, if this extra cost is not offset by the reduction of its own capital cost, the deal is not acceptable. My method treats the ceding insurer and the reinsurer equally. The optimal treaty is fair to both firms and is the most beneficial to the primary policyholder. Obviously, an optimal reinsurance treaty so defined cannot be calculated by either company since one company cannot model the other company's aggregate loss distribution. Fortunately, it is not necessary to explicitly calculate the optimal treaty terms. As long as each company correctly prices its own policies, the optimal treaty is automatically attained through price competition. I will use a few examples to illustrate the working of this market force.

Let us begin with a simple scenario. Assume a primary insurer has written a line of business and would like to cede a part of it. Denote by f_{Pr} the amount of capital cost saved by reinsurance. The reinsurer incurs extra capital costs associated with the assumed loss. It charges the primary insurer an additional premium, denoted by f_{Re}, to cover these costs.[5] So the primary insurer pays an amount of premium f_{Re} to save an amount of cost f_{Pr}. The reinsurance only makes sense if $f_{Re} \leq f_{Pr}$, which means the sum of the capital costs of both companies must decrease.

Assume further that there are two competing reinsurers; a treaty placed with reinsurer 1 costs the primary insurer a premium $f_{Re,1}$ to save a capital cost $f_{Pr,1}$, and one placed with reinsurer 2 costs $f_{Re,2}$ to save a capital cost $f_{Pr,2}$. The immediate (present value) benefits from the treaties are $f_{Pr,1} - f_{Re,1}$ and $f_{Pr,2} - f_{Re,2}$, respectively. The insurer would choose the reinsurer with the greater benefit,

[5] Rigorously, f_{Pr} and f_{Re} represent risk-adjusted present values of the corresponding capital cost cash flows.

which is the one producing the lower total capital cost.

Now look at an example where primary insurers choose reinsurance to compete with each other for business. Suppose that a line of business is on the market and two insurers are bidding. Suppose each insurer has a set of available reinsurance options. As proved in Section 2, the fair gross premium includes a capital cost component that equals the present value of the total capital cost of the insurer and the reinsurer. To win the bid, an insurer looks for a reinsurance treaty that can produce the lowest possible total capital cost. Eventually, the business will go to the insurer able to secure a reinsurance with so low a total capital cost that the other cannot match. Obviously, an insurer's ability to get a more competitive reinsurance deal depends on its existing business and capital structure.

The above analysis shows that market competition always favors a reinsurance structure that produces less total capital cost. Consequently, a reinsurance structure with the least total capital cost is an attracting equilibrium.

4 Capital Requirement Defined by a Coherent Risk Measure

Suppose a uniform capital requirement is imposed on all insurers by regulation. I will only deal with the loss risk, that is, the risk that L becomes very large. The required capital can be defined by a risk measure on the loss distribution. A class of risk measures considered desirable are the coherent risk measures. According to Artzner et al. (1999), risk measure ρ is called coherent if it satisfies the following conditions:

- Monotonicity: For any two losses, L_1 and L_2, if $L_1 \leq L_2$, then $\rho(L_1) \leq \rho(L_2)$
- Positive homogeneity: For any loss L and a constant $a > 0$, $\rho(aL) = a\rho(L)$
- Translation invariance: For any loss L and a constant b, $\rho(L+b) = \rho(L) + b$
- Subadditivity: For any two losses, L_1 and L_2, $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$

All these properties have simple intuitive meanings. Most important to this study is subadditivity. Subadditivity implies diversification: When two risks are pooled together, the required capital of the pool is less than the sum of the required capitals of each risk.

A typical property/casualty loss is a continuous random variable, that is, its cumulative distribution function $F_L(x)$ is continuous. The p-quantile of L is defined by

$$Q_p(L) = \min\{x|F_L(x) \geq p\}, \qquad p \in (0,1), \tag{4.1}$$

and the tail value at risk (TVaR) at level p is

$$\text{TVaR}_p(L) = E[L|L \geq Q_p(L)], \qquad p \in (0,1). \tag{4.2}$$

The TVaR is the most well-known coherent risk measure for continuous risks. (The quantile, also called the value at risk, does not always respect subadditivity.) The TVaR will be used in my illustrative examples.

Suppose a coherent risk measure ρ is selected by the regulator. Then $\rho(L)$ is the amount of assets a company is required to hold. In a one-year model, the premium provides part of the assets at the beginning of the year; the required capital thus equals the required assets minus the premium. Following Section 2, I examine reinsurance structures that minimize the sum of the required capitals of the insurer and the reinsurer. This is equivalent to the problem of minimizing the sum of their required assets,[6] i.e., minimizing the sum of their risk measures. Note that the required assets should be calculated from the loss distribution at the end of the year and discounted back to the beginning of the year. I ignore the discounting here for simplicity.

5 Lower Bound of Total Capitals and Comonotonicity

Reconsider the simplified model with a single loss L, one primary insurer and one reinsurer. The primary insurer issues a policy to cover the entire loss L and cedes part of it to the reinsurer. Thus, L is split between the two insurers, $L = L_{Pr} + L_{Re}$. For a given coherent risk measure ρ, by the rule of subadditivity, $\rho(L) \leq \rho(L_{Pr}) + \rho(L_{Re})$. This inequality provides an absolute lower bound for the sum of capitals: however L is split between the two insurers, the sum of their required assets is no less than $\rho(L)$. To minimize the total required capital is to get the sum $\rho(L_{Pr}) + \rho(L_{Re})$ as close to $\rho(L)$ as possible.

The lower bound can be attained by many reinsurance arrangements. One trivial case is that $L_{Pr} = L$ and $L_{Re} = 0$, or $L_{Pr} = 0$ and $L_{Re} = L$, that is, only one insurer holds all of L. This fact is no surprise, for if there is only one insurer

[6] This can be explained using equations (2.2) and (2.3). The assets for the insurer are $p_n + e_{Pr}$, and that for the reinsurer are $p_c + e_{Re}$. It is proved in Section 2 that the total (gross) fair premium $p_n + p_c$ decreases as the total capital $e_{Pr} + e_{Re}$ decreases. If a reinsurance treaty minimizes the total required assets, it must simultaneously minimize the total required capital and the total fair premium.

and all losses are insured with it, the effect of diversification is maximized, and the least amount of capital is required. An extension of this fact is that an insurance market with few insurers requires less total amount of capital than a market with many insurers. But few insurers means less competition, and insurers have less incentive to price policies fairly.

The lower bound is also reached by the quota share reinsurance. If a is the quota share ceding fraction ($0 < a < 1$), then $L_{Pr} = (1-a)L$ and $L_{Re} = aL$. The equality $\rho(L) = \rho(L_{Pr}) + \rho(L_{Re})$ follows from the rule of positive homogeneity of ρ. More generally, if two losses L_1 and L_2 are perfectly linearly correlated, that is, their linear (Pearson) correlation coefficient equals 1, then $\rho(L_1 + L_2) = \rho(L_1)+\rho(L_2)$. Therefore, if a reinsurance treaty splits L into two linearly correlated parts, then the sum of their required capitals is minimized. The condition of perfect linear correlation can rarely be fulfilled. Fortunately, it can be much relaxed in the following two steps. First, although some kind of perfect correlation has to exist between two losses, L_1 and L_2, for their risk measures to add up, the correlation does not have to be linear—any monotonic and increasing relationship suffices. Second, a perfect correlation only needs to exist at the tail, for large values of L_1 and L_2. Mathematically, both these issues have been well treated in the literature, as explained below.

Let me first give the definition of comonotonicity. Two random variables, X and Y, are perfectly linearly correlated (the linear correlation coefficient of X and Y equals 1) if and only if their support lies in a straight line with a positive slope. (Recall that the support is the set of all possible values of X and Y in the (x, y)-plane. It can be visualized by drawing a scatter plot. A scatter plot of a pair of random variables is merely a small, random subset of its support.) Comonotonicity is an extention of perfect linear correlation. Two random variables, X and Y, are called comonotonic, if their support lies in a one-dimensional curve that is never decreasing. More precisely, the support of a pair of comonotonic random variables satisfies the following condition: if, for any two points in the support, (x_1, y_1) and (x_2, y_2), $x_1 < x_2$ implies $y_1 \leq y_2$ and $y_1 < y_2$ implies $x_1 \leq x_2$. A good overview of comonotonicity and its application in risk theory is Dhaene et al. (2006), where comonotonicity is defined for any number of random variables. Comonotonicity can be considered a perfect nonlinear correlation. For example, if X is a positive random variable, then X and X^2 are comonotonic but not linearly correlated. The support of (X, X^2) is contained in the graph of parabola $y = x^2$. The Spearman rank correlation coefficient is a more meaningful measure than the linear correlation coefficient for characterizing such a nonlinear relationship. The rank correlation coefficient of two comonotonic random variables equals 1 (see Wang 1998), while their linear correlation coefficient is typically less than 1.

The TVaR is a coherent risk measure and is also additive for comonotonic risks: If two losses L_1 and L_2 are comonotonic, then $\mathrm{TVaR}_p(L_1 + L_2) = \mathrm{TVaR}_p(L_1) + \mathrm{TVaR}_p(L_2)$ for any p (Dhaene et al. 2006).[7] In the one-insurer-one-reinsurer model, assume the required asset is determined by a risk measure ρ that is coherent and additive for comonotonic risks. If L is split in such a way that L_{Pr} and L_{Re} are comonotonic, then $\rho(L_{Pr}) + \rho(L_{Re})$ reaches its lower bound $\rho(L)$. We have seen that the quota share reinsurance splits the loss this way. Another example is the stop-loss reinsurance, which is defined by

$$L_{Pr} = \min(L, k), \qquad L_{Re} = \max(L - k, 0), \tag{5.1}$$

where $k > 0$ is the attachment point. It is easy to check that the three variables L, L_{Pr} and L_{Re} are comonotonic, and $\rho(L) = \rho(L_{Pr}) + \rho(L_{Re})$.

Risk measures like $Q_p(L)$ and $\mathrm{TVaR}_p(L)$ are determined by large values of L. When looking for a way to split L into L_{Pr} and L_{Re} to minimize the total capital, one should focus on large losses. The condition of comonotonicity requires the entire support of the random vector to be in a one-dimensional non-decreasing curve. This condition is too strong. Cheung (2009) introduces the concept of upper comonotonicity, only requiring the condition to be satisfied in the upper tail. If L_{Pr} and L_{Re} are upper comonotonic, then $\rho(L) = \rho(L_{Pr}) + \rho(L_{Re})$, where ρ is either Q_p or TVaR_p and p is sufficiently close to 1. In general, the amount of total capital corresponding to a reinsurance structure is determined by large losses only.

6 Optimal Reinsurance in a General Setting

I now apply the concepts developed so far to formulate a general problem about optimal reinsurance. I have discussed the problem of splitting a single loss L between an insurer and a reinsurer. In the real world, a primary insurer does not have the option or the intension to cover its entire book with a reinsurance treaty. It only attempts to cede some unwanted lines or accounts. On the other hand, a reinsurer assumes losses from many insurers and reinsurers. A new treaty adds losses to its existing book. When determining the optimal reinsurance, one needs

[7] There are other risk measures that are coherent and additive for comonotonic risks, e.g., the concave distortion risk measures. The VaR is additive for comonotonic risks but is not coherent (Dhaene et al. 2006).

to consider these "other" loss portfolios of the ceding insurer and the reinsurer, in addition to the loss portfolio to be ceded. The following model includes all these sets of losses.

Assume a primary insurer initially carries losses $X + Z$, where X will be entirely retained and Z may be partially ceded. The reinsurer holds a loss Y before assuming any part of Z. A reinsurance treaty splits Z into a net and a ceded part, $Z = Z_n + Z_c$. Before reinsurance, the total required asset of the insurer and the reinsurer is $\rho(X + Z) + \rho(Y)$. After reinsurance, the total required asset is $\rho(X + Z_n) + \rho(Y + Z_c)$. A treaty is optimal if the latter sum is minimized.

If ρ is a coherent risk measure, an absolute lower bound for $\rho(X + Z_n) + \rho(Y + Z_c)$ is $\rho(X + Y + Z)$. In general, the distributions of losses X, Y and Z and correlations between them are complex. There is no ceding arrangement that can bring down the sum $\rho(X + Z_n) + \rho(Y + Z_c)$ to anywhere near this lower bound. Moreover, in the reinsurance market, only a few types of treaties are commonly placed, like the quota share, excess of loss, catastrophe and stop loss treaties. This further limits how low $\rho(X + Z_n) + \rho(Y + Z_c)$ can become. Minimizing the sum $\rho(X + Z_n) + \rho(Y + Z_c)$ for a given set of available treaties is mathematically a constrained optimization problem.

From the preceding section, we learned that if a ceding arrangement makes $X + Z_n$ and $Y + Z_c$ comonotonic (upper comonotonicity suffices), then the sum of required capitals attains its minimum value $\rho(X + Y + Z)$. In other words, the minimum sum of capitals corresponds to the maximum correlation between the losses (their rank correlation equals 1). This suggests that the value of $\rho(X + Z_n) + \rho(Y + Z_c)$ may be inversely related to the correlation between $X + Z_n$ and $Y + Z_c$. A reinsurance contract that makes the total capital small must make the correlation large. This observation, if it can be proved in certain circumstances, should be very interesting. I will examine some examples where a linkage between the total capital and the correlation does exist. In the appendix, I will provide a graphic reasoning to further support this relationship.

7 Examples

In the rest of the paper, examples are provided to examine how closely the total capital is related to the correlation between the ceding insurer's and the reinsurer's losses. The first example uses normally distributed losses, where the optimal ceding terms can be obtained in closed form. The second example is more general and has to be solved numerically. The optimal cedings are calculated based on simulation results.

7.1 A multivariate normal example

Let X, Y and Z be three jointly normally distributed variables. X and Z are losses written by the primary insurer, X will be retained and Z partially ceded; Y is the existing loss of the reinsurer. Suppose only quota share treaties may be placed on Z. Although this is not a realistic situation (actual losses do not take negative values as the normal distribution does), discussion of this tractable example can provide us valuable insights.

Let X, Y and Z have the following parameters: means μ_x, μ_y and μ_z, standard deviations σ_x, σ_y and σ_z, and pairwise correlation coefficients γ_{xz}, γ_{yz} and γ_{xy}. If a quota share treaty is placed and a is the ceding fraction, then the primary company's net loss is $L_{Pr} = X + (1-a)Z$, and the reinsurer's total loss is $L_{Re} = Y + aZ$. These two losses are also normal random variables. Their means and standard deviations are as follows.

$$\begin{aligned}
\mu_{Pr} &= E(L_{Pr}) = \mu_x + (1-a)\mu_z \\
\sigma^2_{Pr} &= \mathrm{Var}(L_{Pr}) = \sigma^2_x + (1-a)^2\sigma^2_z + 2(1-a)\gamma_{xz}\sigma_x\sigma_z \\
\mu_{Re} &= E(L_{Re}) = \mu_y + a\mu_z \\
\sigma^2_{Re} &= \mathrm{Var}(L_{Re}) = \sigma^2_y + a^2\sigma^2_z + 2a\gamma_{yz}\sigma_y\sigma_z
\end{aligned}$$

For a given confidence level p, the risk measures Q_p and TVaR_p of a normal random variable can be easily obtained. In fact, they can be written as $Q_p = \mu + h_p\sigma$ and $\mathrm{TVaR}_p = \mu + k_p\sigma$, where h_p and k_p are constants independent of μ and σ. For example, $Q_{0.99} = \mu + 2.33\sigma$ and $\mathrm{TVaR}_{0.99} = \mu + 2.67\sigma$. Therefore, if the risk measure ρ is of the quantile or the TVaR type, minimizing the sum $\rho(L_{Pr})+\rho(L_{Re})$ is equivalent to minimizing the sum $\sigma_{Pr} + \sigma_{Re}$. The latter problem will be solved below.

The variances of the insurer and the reinsurer can be written in a simpler form

$$\begin{aligned}
\sigma^2_{Pr} &= \sigma^2_z((a - A_{Pr})^2 + B^2_{Pr}) \\
\sigma^2_{Re} &= \sigma^2_z((a + A_{Re})^2 + B^2_{Re}),
\end{aligned} \tag{7.1}$$

where

$$\begin{aligned}
A_{Pr} &= 1 + \gamma_{xz}\sigma_x/\sigma_z, \qquad B^2_{Pr} = (1-\gamma^2_{xz})\sigma^2_x/\sigma^2_z \\
A_{Re} &= \gamma_{yz}\sigma_y/\sigma_z, \qquad B^2_{Re} = (1-\gamma^2_{yz})\sigma^2_y/\sigma^2_z.
\end{aligned} \tag{7.2}$$

The sum of standard deviations is thus

$$\sigma_{Pr} + \sigma_{Re} = \sigma_z\left(((a - A_{Pr})^2 + B^2_{Pr})^{1/2} + ((a + A_{Re})^2 + B^2_{Re})^{1/2}\right).$$

To minimize this sum is to minimize the following function $f(a)$

$$f(a) = ((a - A_{Pr})^2 + B_{Pr}^2)^{1/2} + ((a + A_{Re})^2 + B_{Re}^2)^{1/2},$$

where the ceding fraction a is between 0 and 1. The derivative of $f(a)$ is

$$f'(a) = \frac{a - A_{Pr}}{((a - A_{Pr})^2 + B_{Pr}^2)^{1/2}} + \frac{a + A_{Re}}{((a + A_{Re})^2 + B_{Re}^2)^{1/2}}.$$

Setting the right-hand side of the equation equal to zero, moving one of the terms to the other side and squaring the terms, we have

$$\frac{(A_{Pr} - a)^2}{(a - A_{Pr})^2 + B_{Pr}^2} = \frac{(a + A_{Re})^2}{(a + A_{Re})^2 + B_{Re}^2}.$$

Simplifying this gives

$$(A_{Pr} - a)^2 B_{Re}^2 = (a + A_{Re})^2 B_{Pr}^2.$$

Let us assume that $\gamma_{xz} \geq 0$ and $\gamma_{yz} \geq 0$, meaning that the losses X, Y and Z are not negatively correlated, a condition likely to be true in the real world. Mathematically, this implies $A_{Pr} \geq 1$ and $A_{Re} \geq 0$. If we assume $-A_{Re} \leq a \leq A_{Pr}$, then $A_{Pr} - a \geq 0$ and $a + A_{Re} \geq 0$. Taking the square root in the above equation, we get the solution

$$a^* = \frac{A_{Pr} B_{Re} - A_{Re} B_{Pr}}{B_{Pr} + B_{Re}}. \tag{7.3}$$

This is the unique zero of $f'(a)$ between $-A_{Re}$ and A_{Pr} and the unique minimum point of $f(a)$. The function $f(a)$ strictly decreases from $-A_{Re}$ to a^* and strictly increases from a^* to A_{Pr}. Note that the optimal ceding fraction does not depend on how X and Y are correlated.

Now let us examine a few special cases. First, suppose Z is uncorrelated with both X and Y, that is, $\gamma_{xz} = \gamma_{yz} = 0$. From the equations (7.2), $A_{Pr} = 1$, $B_{Pr} = \sigma_x/\sigma_z$, $A_{Re} = 0$ and $B_{Re} = \sigma_y/\sigma_z$. Using (7.3), we obtain the optimal ceding fraction $a^* = \sigma_y/(\sigma_x + \sigma_y)$. So, in this case, to minimize $\sigma_{Pr} + \sigma_{Re}$, Z should be shared between the primary insurer and the reinsurer in proportion to the standard deviations of their "fixed" losses, σ_x and σ_y.

A more interesting case is when Z is highly correlated to X but almost uncorrelated to Y. Then $\gamma_{xz} \approx 1$ and $\gamma_{yz} \approx 0$. These imply that $A_{Pr} \approx 1 + \sigma_y/\sigma_z$, $B_{Pr} \approx 0$, $A_{Re} \approx 0$ and $B_{Re} \approx \sigma_y/\sigma_z$. By (7.3), $a^* \approx 1 + \sigma_x/\sigma_z$. This a^* is greater than 1. Thus, to minimize $\sigma_{Pr} + \sigma_{Re}$, Z should be 100 percent ceded. On the other hand, since Z and X are highly correlated, the more Z is ceded to the reinsurer, the greater is the (linear) correlation between $X+(1-a)Z$ and $Y+aZ$. This correlation is maximized at $a = 100\%$. In this example, the reinsurance is optimized at the same ceded ratio where the correlation between the losses is maximized.

A parallel result is that, if Z is highly correlated to Y but almost uncorrelated to X, then the optimal ceded ratio is 0 percent. At this ceded ratio, the correlation between the losses is again maximized.

Now let us plug in some numerical values. Assume $\sigma_x = 300$, $\sigma_y = 500$ and $\sigma_z = 100$; $\gamma_{xz} = 0.4$, $\gamma_{yz} = 0.4$ and $\gamma_{xy} = 0.2$. Using (7.2), we compute $A_{Pr} = 2.20$, $B_{Pr} = 2.75$, $A_{Re} = 2.00$ and $B_{Re} = 4.58$. Substituting these into (7.3), we obtain the optimal ceding fraction $a^* = 62.5\%$. However, this a^* does not provide the maximum correlation between $X + (1 - a)Z$ and $Y + aZ$. Using simulation, we get that the maximum linear correlation coefficient is 0.290 and is reached at the ceded ratio of 30.5 percent. Therefore, the minimum total capital does not always correspond to the maximum correlation. As mentioned before, this result is not really a surprise because the capital is determined by large losses, while the linear or rank correlation coefficient does not distinguish between large and small losses (or even negative losses, which is the case in this example).

7.2 A numerical example

If the joint distribution of losses X, Y and Z is known, and a set of available reinsurance treaties is given, the optimal treaty can be found by simulation. To have an easy control on correlations between the losses, I will assume the losses are jointly lognormal. I will look at two common types of treaties, the quota share and the stop loss.

Let the variables X, Y and Z be jointly lognormal, in the sense that $\ln(X)$, $\ln(Y)$ and $\ln(Z)$ are jointly normal. The mean μ^0 and the standard deviation σ^0 of these normal variables are as follows.

	$\ln(X)$	$\ln(Y)$	$\ln(Z)$
μ^0	19.5	20.0	17.0
σ^0	0.16	0.25	1.10

The mean, the standard deviation and quantiles of X, Y and Z can be computed from the above table with simple formulas. I will denote a parameter for a normal random variable with a superscript 0, and the same parameter for the corresponding lognormal variable without a superscript. For example, μ_x^0 is the mean of $\ln(X)$ and μ_x the mean of X. These formulas are well known:

$\mu_x = \exp(\mu_x^0 + (\sigma_x^0)^2/2)$, and $\sigma_x = \exp(\mu_x^0 + (\sigma_x^0)^2/2)(\exp((\sigma_x^0)^2) - 1)^{1/2}$. The p-quantile of X can be written as $Q_p(X) = \exp(\mu_x^0 + h_p \sigma_x^0)$, where h_p is the p-quantile of the standard normal distribution. More complex measures of the lognormals, like $\text{TVaR}_p(X)$ or the standard deviation of $X + Y + Z$, are more conveniently estimated using simulation. Some useful statistics for X, Y and Z are listed in the following table (loss amounts are in millions).

	X	Y	Z
μ	298	501	44
σ	48	127	68
CV	0.16	0.25	1.53
$Q_{0.99}$	427	868	312
$\text{TVaR}_{0.99}$	451	941	477

I will choose $\rho = \text{TVaR}_{0.99}$ as the risk measure. In addition to the known μ and σ, if the linear correlation coefficients γ_{xz}, γ_{yz} and γ_{xy} are also given, then the distribution of the triplet (X, Y, Z) is completely determined. Following our naming convention, γ_{xz}^0 is the linear correlation coefficient between $\ln(X)$ and $\ln(Z)$. γ_{xz}^0 determines γ_{xz}, and vise versa. A greater γ_{xz}^0 corresponds to a greater γ_{xz}. The strongest correlation between X and Z is attained when $\ln(X)$ is a linear function of $\ln(Z)$ with a positive slope. In this case, $\gamma_{xz}^0 = 1$ but γ_{xz} is generally less than 1.[8]

A straightforward sampling method is used to find the optimal ceding term. For μ and σ in the above table and known γ_{xz}, γ_{yz} and γ_{xy}, a large random sample of (X, Y, Z) is drawn (using Excel with the @RISK add-in or with a macro performing the Cholesky decomposition). Applying a given reinsurance treaty on the sample data, we get samples of losses of the primary insurer and the reinsurer, from which the TVaR of the losses can be estimated. Table 1 displays results for quota share treaties. Five scenarios of different γ_{xz}^0, γ_{yz}^0 and γ_{xy}^0 are analyzed. For each scenario, a set of 20,000 sample points of the triplet (X, Y, Z) is drawn; 101 quota share fractions, a, ranging from 0 to 100 percent with 1 percent increments, are applied; the measures $\rho(X + (1 - a)Z)$ and $\rho(Y + aZ)$ are estimated; and the least sum of them is found by comparison, which gives the optimal quota share term. (Loss amounts in Table 1 are in millions.)

[8] The exact formula is $\gamma_{xz} = [\exp(\sigma_x^0 \sigma_z^0 \gamma_{xz}^0) - 1]/[(\exp((\sigma_x^0)^2) - 1)(\exp((\sigma_z^0)^2) - 1)]^{1/2}$. When $\gamma_{xz}^0 = 1$, γ_{xz} is generally less than 1, but the Spearman rank correlation coefficient between X and Z equals 1.

Table 1: Optimal Quota Share Fractions

	(1)	(2)	(3)	(4)	(5)
γ^0_{xz}	0.9	0.9	0	0.1	0
γ^0_{yz}	0	0.1	0.99	0.9	0
γ^0_{xy}	0	0	0	0	0
$\rho(X+Y+Z)$	1,540	1,570	1,764	1,721	1,422
a^* (optimal ceding)	100%	100%	0%	36%	75%
$\rho(X+(1-a^*)Z)+\rho(Y+a^*z)$	1,596	1,624	1,771	1,756	1,529

In the table, $\rho(X+Y+Z)$ is the absolute lower bound of the total required asset, for any type of reinsurance. In scenario (3), the optimal total asset $\rho(X+(1-a^*)Z)+\rho(Y+a^*Z)$ is close to $\rho(X+Y+Z)$. But, in general, the difference between the two is sizable. In scenarios (1) and (2), Z is strongly correlated to X but weakly correlated Y. Ceding out the entire Z ($a=100\%$) would maximize the correlation between $X+(1-a)Z$ and $Y+aZ$.[9] This supports the claim that the optimal treaty is the one that creates the strongest correlation between the insurer's and the reinsurer's losses. A similar relationship holds in scenario (3), where Z is strongly correlated to Y but weakly correlated to X. The optimal term is to cede nothing, which again corresponds to the strongest correlation between the two losses. However, in scenario (5), the optimal ceding ratio is 75 percent, while, as can be shown, the maximum correlation is reached at $a=55\%$. The two ratios are different.

I now consider the same five correlation scenarios and perform a similar analysis for stop-loss treaties. In each scenario, let the primary insurer's retention, k, vary from 20 million to 250 million, with 5 million increments. The ceded loss is $Z_c=\max(Z-k,0)$, and the retained loss $Z_n=Z-Z_c=\min(Z,k)$. Comparing the total asset $\rho(X+Z_n)+\rho(Y+Z_c)$ for all these k, we get the optimal retention k^*. The results are summarized in Table 2 (loss amounts are in millions).

In the first two scenarios, Z is highly correlated to X; in the next two scenarios, it is highly correlated to Y. Thus, intuitively, in the first two scenarios, the correlation (at the right tail) between $X+Z_n$ and $Y+Z_c$ increases as more of Z is ceded. In fact, the sample linear correlation is indeed the highest at $k=20$.

[9] It can be proved mathematically that, if γ_{xz} is very close to 1, then the greater the ceded ratio a, the greater the linear correlation between $X+(1-a)Z$ and $Y+aZ$. The intuition behind this result is that, if Z behaves very similarly to X, then $Y+Z$, for an arbitrary variable Y, behaves more similarly to X than Y does.

Table 2: Optimal Stop-Loss Retentions

	(1)	(2)	(3)	(4)	(5)
γ^0_{xz}	0.9	0.9	0	0.1	0
γ^0_{yz}	0	0.1	0.99	0.9	0
γ^0_{xy}	0	0	0	0	0
$\rho(X+Y+Z)$	1,540	1,570	1,764	1,721	1,422
k^* (optimal retention)	20	20	250	250	85
$\rho(X+Z^*_n)+\rho(Y+Z^*_c)$	1,598	1,625	1,804	1,770	1,557

$Z^*_n = \min(Z, k^*)$, $Z^*_c = \max(Z - k^*, 0)$

This again supports the claim that the optimal treaty maximizes the correlation. This statement holds true in the next two scenarios, where the optimal treaty is to cede the least of Z. However, in scenario (5), the maximum linear correlation is attained at the retention $k = 115$, which is different from the optimal retention $k^* = 85$.

Finally, let us look at scenario (5) and compare the two types of treaties. The optimal total required asset for the stop-loss treaties is 1,557, and for the quota share treaties it is 1,529. So the quota share is more effective in cutting the total capital.[10] This appears to contradict the general belief that a stop-loss treaty reduces volatility more effectively than a quota share treaty. The fact is, however, although the stop-loss treaty cuts more capital from the primary insurer, it adds even more to the reinsurer, which results in an increase in the total required capital. In general, which type of treaty reduces the total capital more effectively depends on the joint distribution of all losses.

8 Conclusions

I have proposed to call a reinsurance arrangement optimal if it minimizes the total capital of the primary insurer and the reinsurer. This optimal reinsurance produces the lowest price for primary insurance policies, so is an attracting equilibrium under market competition. An interesting relationship is observed between the total capital and the tail correlation between the losses of the insurer and the

[10] The quota share structure is better in the other four scenarios as well, but those results are of no surprise. As the stop-loss retention is limited to between 20 and 250, ceding the whole of Z and ceding none of Z are excluded, yet the optimal quota share terms in these scenarios fall into these extremes.

reinsurer. A multivariate normal model and a numerical example are analyzed to get more insight into the nature of an optimal treaty.

This paper fills a gap in the existing literature on optimal reinsurance, in which the capital cost of the reinsurer has not been adequately addressed. My approach establishes a close link between reinsurance and pricing of insurance and reinsurance policies. In a competitive market, reinsurance not only provides the ceding insurer a tool of risk transfer, but also satisfies the reinsurer with a fair amount of profit and benefits primary policyholders by reducing their costs.

Tail correlation between losses has been widely discussed in relation to risk measurement and management. In this paper, it is linked to the size of the total capital. This seems to be an interesting area of research.

References

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. "Coherent Measures of Risk." *Mathematical Finance* 9: (1999) 203–228.

Cai, Jun, and Ken Seng Tan. "Optimal Retention for a Stop-Loss Reinsurance Under the VaR and CTE Risk Measures." *ASTIN Bulletin* 37, 1: (2007) 93–112.

Chandra, Victor, and Michael Sherris. "Capital Management and Frictional Costs in Insurance." *Australian Actuarial Journal* 12, 4: (2006) 399–448.

Cheung, Ka Chun. "Upper Comonotonicity." *Insurance: Mathematics and Economics* 45, 1: (2009) 35–40.

Cummins, J. David. "Asset Pricing Models for Insurance Ratemaking." *ASTIN Bulletin* 20, 2: (1990) 125–166.

Dhaene, Jan, Steven Vanduffel, M. J. Goovaerts, Rob Kaas, Qihe Tang, and David Vyncke. "Risk Measures and Comonotonicity: A Review." *Stochastic Models* 22: (2006) 573–606.

Fu, Luyang, and C.K. "Stan" Khury. "Optimal Layers for Catastrophe Reinsurance." *Variance* 4, 2: (2010) 191–208.

Gajek, Leslaw, and Dariusz Zagrodny. "Reinsurance Arrangements Maximizing Insurer's Survival Probability." *Journal of Risk and Insurance* 71, 3: (2004) 421–435.

Jensen, Michael, and William Meckling. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3, 4: (1976) 305–340.

Kaluszka, Marek. "Optimal Reinsurance Under Mean-Variance Premium Principles." *Insurance: Mathematics and Economics* 28: (2001) 61–67.

Lampaert, Ilse, and Jean-Francois Walhin. "On the Optimality of Proportional Reinsurance." *Casualty Actuarial Society Forum* 2005, spring: (2005) 93–114.

Myers, Stewart S., and Richard A. Cohn. "A Discounted Cash Flow Approach to Property-Liability Insurance Rate Regulation." In *Fair Rates of Return in Property-Liability Insurance*, edited by J. David Cummins, and Scott E. Harrington. Dordrecht: Kluwer Nijhoff Publishing, 1987, 55–77.

Perold, Andre. "Capital Allocation in Financial Firms." *Journal of Applied Corporate Finance* 17, 3: (2005) 110–118.

Taylor, Greg C. "Fair Premium Rating Methods and the Relations Between Them." *Journal of Risk and Insurance* 61: (1994) 592–615.

Vaughn, Trent. "The Impact of Investment Strategy on the Market Value and Pricing Decisions of a Property/Casualty Insurer." *Proceedings of the Casualty Actuarial Society* LXXXV: (1998) 211–244.

Wang, Shaun. "Aggregation of Correlated Risk Portfolios: Models and Algorithms." *Proceedings of the Casualty Actuarial Society* LXXXV: (1998) 848–939.

Zhang, Yingjie. "Economic Impact of Capital Level in an Insurance Company." *Variance* 2, 1: (2008) 39–51.

Appendix. More on the Linkage Between the Total Capital and Correlation

I have shown that the TVaR is a subadditive risk measure: If $\rho = \text{TVaR}_p$, then $\rho(X) + \rho(Y) \geq \rho(X+Y)$, and the equality holds if X and Y (representing the losses of a primary insurer and a reinsurer) are comonotonic. Following this fact, I propose that a linkage exists between the total asset $\rho(X) + \rho(Y)$ and the correlation between X and Y, that is, the greater the tail correlation, the closer is $\rho(X) + \rho(Y)$ to $\rho(X+Y)$. In this appendix, I will use the scatter plot to further explain why there should be such a link.

Figures 1 through 3 provide scatter plots of a pair of losses X and Y correponding to three different correlation scenarios. (The correlations are actually only different at the right tail.) Each loss is in the range $[0, 100)$. In Figure 1, X and Y are comonotonic at the tail. In Figure 2, they are not comonotonic but are still highly correlated at the tail: as X moves up from about 80, Y generally moves up as well, although it sometimes moves in the opposite direction (down) slightly. In Figure 3, X and Y have little correlation at the tail.

Let the risk measure be $\rho = \text{TVaR}_{0.9}$. There are 100 sample points in each figure. The point labeled A has the 11th largest x coordinate, and the one labeled B has the 11th largest y coordinate. The quantile $Q_{0.99}(X)$ is the x coordinate of A, and $Q_{0.99}(Y)$ the y coordinate of B. $\rho(X)$ is the average of the x coordinates of the points to the right of A, and $\rho(Y)$ the average of the y coordinates of the points higher than B. $\rho(X+Y)$ is the average of the largest 10 $x+y$ of all points.

In Figure 1, A and B are actually the same point $(78, 76)$ (coordinates are rounded), and the points to the right of A are the same as those higher than A, which are also the 10 points with the largest $x+y$. Thus, $Q_{0.99}(X) + Q_{0.99}(Y) = Q_{0.99}(X+Y) = 78+76 = 154$, and $\rho(X)+\rho(Y) = \rho(X+Y)$ $(= 178)$. This explains that if X and Y are perfectly correlated at the tail, then $\rho(X)+\rho(Y) = \rho(X+Y)$.

In Figure 2, the upper-right tail is a rather "thin" set. Thus the two points A and B are close to each other. Further, the following three sets of points are similar (contain mostly the same points): those to the right of A, those higher than B, and the ten points with the largest $x+y$. This implies that $\rho(X)+\rho(Y)$ is close to $\rho(X+Y)$. (Here $\rho(X) = 95.3$, $\rho(Y) = 88.8$ and $\rho(X+Y) = 183.9$.) This example shows that if X and Y are highly correlated at the tail, then $\rho(X)+\rho(Y)$ is (greater than but) close to $\rho(X+Y)$.

The upper-right tail in Figure 3 is not a thin set, and the two points A and B are generally far apart. Also, it is likely that the three sets—the one to the right of A, the one higher than B and the one with the largest $x+y$—contain very different points. So $\rho(X)+\rho(Y)$ can be much larger than $\rho(X+Y)$. (Here $\rho(X) = 95.3$, $\rho(Y) = 82.4$ and $\rho(X+Y) = 174.1$.) This is what normally happens when X and Y are not correlated at the tail.

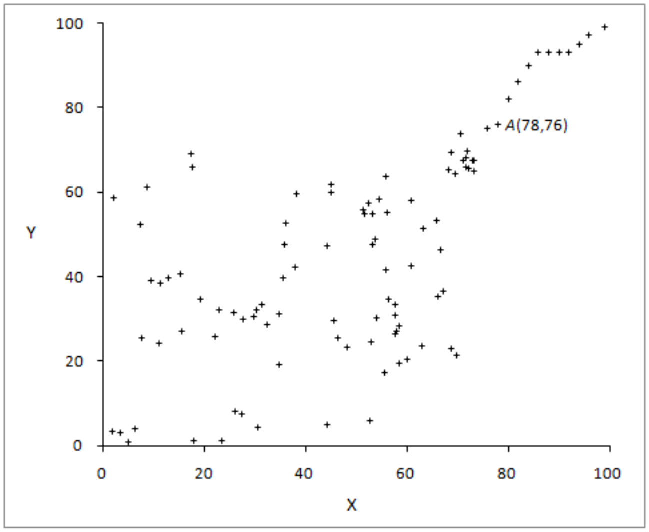


Figure 1: X and Y are comonotonic at the tail; $\rho(X) + \rho(Y) = \rho(X + Y)$

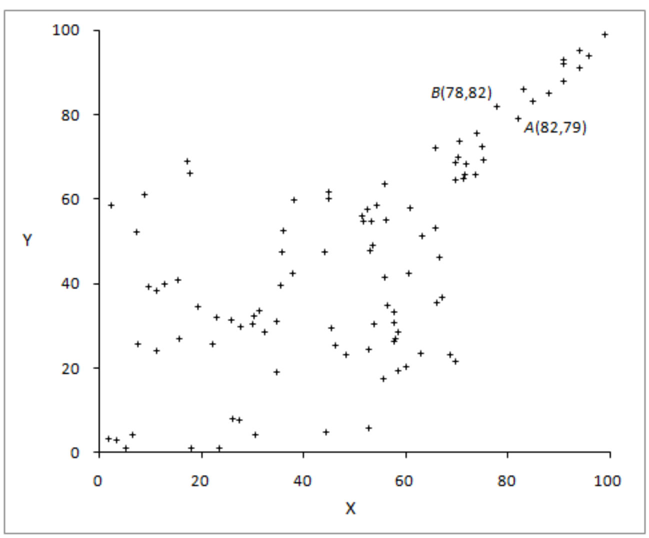


Figure 2: X and Y are highly correlated at the tail; $\rho(X) + \rho(Y)$ is close to (but greater than) $\rho(X + Y)$

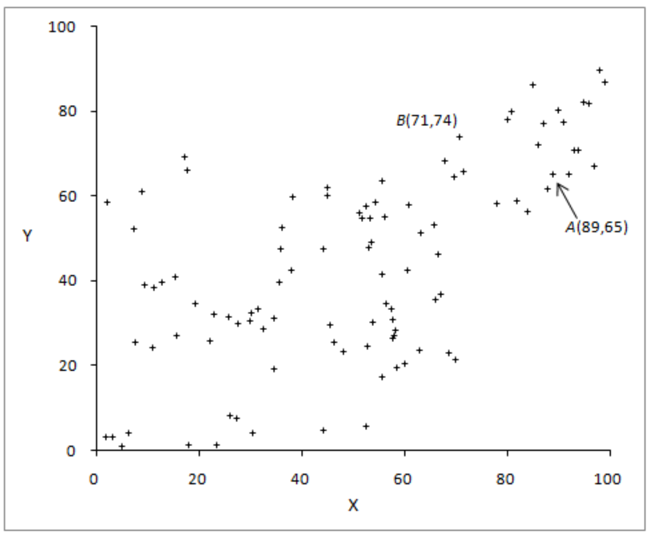


Figure 3: X and Y are not correlated at the tail; $\rho(X) + \rho(Y)$ is generally much greater than $\rho(X + Y)$