Casualty Actuarial Society
E-Forum, Spring 2021

The CAS *E-Forum*, Spring 2021

The Spring 2021 edition of the CAS *E-Forum* is a cooperative effort between the CAS *E-Forum* Committee and various CAS committees, task forces, working parties and special interest sections. This *E-Forum* contains papers written in response to a call for essays on the implications of COVID-19 on the property-casualty insurance industry. The E-Forum also includes four independent research papers.

CAS *E-Forum*, Spring 2021

Table of Contents

E-Forum Committee

For information on submitting a paper to the *E-Forum*, visit https://www.casact.org/publications-research/guide-submission-e-forum.

Interplay between Epidemiology and Actuarial Modeling

Runhuan Feng, Ph.D., FSA, CERA; Longhao Jin, MS;
and Sooie-Hoe Loke, Ph.D.

Abstract
In an era where rapidly evolving situations is the new normal, collaboration between multiple disciplines offers a concerted effort and provides a comprehensive perspective in tackling problems and challenges. In this essay, we illustrate various modeling tools and key ideas used in epidemiology that can be applied to the insurance framework. Specifically, we give an overview of the compartmental models, network models, and agent-based models, and discuss their applications to epidemic and cyber insurance coverages.

Keywords. Compartmental model, network models, agent-based models, epidemic, reserve level, cyber insurance

1. INTRODUCTION

The actuarial profession is constantly evolving. The age of big data has prompted actuaries to learn and apply the state-of-the-art predictive analytics methods and machine learning techniques from non-traditional fields. The current COVID-19 pandemic has changed the world in many ways, and in the insurance industry, has prompted us to think about the challenges posed in the profession and ways to address these problems.

Written from an academic perspective, the goal of this essay is to discuss several modeling tools from epidemiology and how they can be applied to an insurance setting. Rather than reinventing the wheel, actuarial researchers and practitioners could take advantage of a wide variety of epidemic models developed in the medical literature to estimate the evolution of mortality and morbidity rates during a pandemic. We believe that many of these models can be easily adapted for actuarial applications. As both actuaries and epidemiologists share a common interest in using advanced mathematical tools to model risk, we hope that this essay motivates actuarial researchers to utilize these models as building blocks to study insurance coverages targeting infectious diseases.

In the following sections, we present three distinct levels of modeling in epidemiology and describe their applications to insurance modeling. We begin with widely used compartmental models, including both deterministic and stochastic versions. This first level of modeling assumes homogeneous mixing, that is, everyone in the population is equally likely to be susceptible and infectious regardless of age, sex, social structure, etc. The next level of modeling is called contact network models, where the individuals are connected by links to describe interactions between them. Finally, we discuss agent-based models which are microscopic models used to simulate real-world complex patterns. Within each model, we provide a list of references in epidemiology as well as in actuarial science.

2. COMPARTMENTAL MODELS

Consider a population of size $N(t)$ which is indexed by time t. The basic framework of compartmental models is to classify the entire population into several distinct compartments or categories. We illustrate this idea via the well-known SIR model based on the seminal work of Kermack and McKendrick (1927). Consider the following three compartments: **S**usceptible, **I**nfected, and **R**emoved. There are many variations of the SIR model, so in what follows we lay out the assumptions for the most basic SIR model. First, there are no births or immigration, which implies that $N(t) = N$, and the population mixes homogeneously. Second, there is an infection rate β and a recovery rate α. In other words, $1/\beta$ is the average time between contacts and $1/\alpha$ is the average time until removal. Finally, once recovered (or removed), an individual is immune and can no longer spread the disease. The following figure summarizes some of the assumptions in the model:

$$\boxed{\text{Susceptible}} \xrightarrow{\beta} \boxed{\text{Infected}} \xrightarrow{\alpha} \boxed{\text{Removed}}$$

This gives rise to the following system of differential equations:

$$S'(t) = -\beta\, S(t)\, I(t) \tag{2.1}$$

$$I'(t) = \beta\, S(t)\, I(t) - \alpha\, I(t) \tag{2.2}$$

where $S(t)$ and $I(t)$ are the number of susceptible and infected individuals, respectively. The number of removed individuals is thus $N - S(t) - I(t)$.

A main concern in epidemiology modeling is whether a disease will spread upon its introduction. Observe from (2.2) that the disease will spread (i.e., $I'(t)$ is positive) if $\beta S(t) - \alpha > 0$ and the disease will die out if $\beta S(t) - \alpha < 0$. This motivates the relevance of a key quantity called the effective (or general) reproductive number, given by $R_t = \beta S(t)/\alpha$, which represents the average number of secondary infections due to a single infectious individual at a given time t. It is also worth noting that β/α is the average number of contacts by an infected person with others before removal. When $t = 0$, R_0 (read as R-naught) is known as the basic reproduction number, which is a common measure used to determine if a disease will spread out during the early phase of the outbreak. If $R_0 > 1$, the disease will start to spread, but not if $R_0 < 1$. Zhao et al. (2020) gave a preliminary estimate of R_0 of the coronavirus to be between 2.24 and 3.58.

More details regarding the SIR model as well as other compartmental models like SIS (Susceptible-Infected-Susceptible) and SEIR (Susceptible-Exposed-Infected-Removed) can be found in classic monographs such as Anderson and May (1992), Brauer and Castillo-Chavez (2012), and Diekmann et al. (2012). For a more recent account of developments in mathematical epidemiology, see Brauer (2017). The aforementioned monographs also discussed heterogeneous mixing models (incorporating age, social structure, etc.) as well as cross-population models.

2.1 Insurance Application

Using the basic SIR model, Feng and Garrido (2011) proposed several insurance policies and models to quantify infection risk using actuarial methodologies. For example, susceptible individuals pay premiums at a constant rate π and once infected, the insurer will pay hospitalization benefits, say at a constant rate of 1. For this particular policy, the insurer's reserve level is

$$V(\pi,t) = \pi \int_0^t s(x)dx - \int_0^t i(x)dx, \tag{2.3}$$

where $s(t) = S(t)/N$ and $i(t) = I(t)/N$. In analyzing the reserve, there are four possible shapes of graph (as a function of time) and it was recently discovered that these shapes are closely linked to the effective reproduction number, R_t, as summarized in the following table:

Shape of $V(\pi,t)$	Interval for values of π
Increasing concave	At least $(1/R_T) - 1$
Increasing concave-then-convex	Between $(1/R_{tm}) - 1$ and $(1/R_T) - 1$
Nonmonotonic concave-then-convex	Between $(1/R_0) - 1$ and $(1/R_{tm}) - 1$
Nonmonotonic convex	Between 0 and $(1/R_0) - 1$

Here, T is the time of the disease-free state (i.e., $I(T) = 0$) and the exact expression of R_{tm} is provided in Feng and Garrido (2011). It is interesting to note that a related quantity from the table, namely $1 - 1/R_0$, is called the herd immunity threshold (see Fine et al. (2011)).

This deterministic insurance model was later extended to a stochastic one by Lefèvre et al. (2017), which provided many elegant results on integral functionals using martingale arguments. The work of Lefèvre et al. (2017) in stochastic SIR insurance model has since opened the door to many classical problems in probability. Lefèvre and Picard (2018a) generalized the SIR model to a controlled epidemic model, where the infectious will be isolated by health organizations to ease the severity of the disease and studied the representation of epidemic outcomes and path integrals in terms of pseudo-polynomials. Lefèvre and Simon (2018) considered cross-infection between two linked populations. A general approach to study Laplace transforms of these integral functionals was developed by Lefèvre and Picard (2018b). More recently, Lefèvre and Simon (2019) proposed a general block-structured Markov processes for epidemic modeling. For a thorough discussion of stochastic epidemic models and methods for their statistical analysis, see Andersson and Britton (2012).

In terms of other deterministic insurance models, Perera (2017) considered the control strategy in the simple SIR model as well as the variation of the premium with respect to the model parameters. Nkeki and Ekhaguere (2020) constructed the SIDRS model and studied its insurance applications. Billard and Dayananda (2014b) and Billard and Dayananda (2014a) developed a multi-stage HIV/AIDS model considering non-disease death in each compartment, where the waiting time distribution is used to measure the total amount of time one individual holds in one state. Not only that premiums are defined by different insurance functions, but health-care cost adjustments are also included. Shemendyuk et al. (2019) investigated the deterministic and stochastic SIR models with multiple centers and migration fluxes. The optimal health-care premium is determined by considering different vaccine allocation strategies. Optimal resource allocation and contingency planning were also addressed in Chen et al. (2020).

3. NETWORK MODELS

To capture the interactions among individuals while still preserving some aspects of the compartmental models, one can use a network model. Using the language of graph theory, the individuals are known as vertices or nodes and are connected by edges or links, which indicate relationships between various vertices. Some common metrics used to describe a network include shortest path length, degree distribution, and clustering coefficient. Epidemiologists have a long history of using network models to study diseases. An extensive survey by Pastor-Satorras et al. (2015) describes various types of network models used in epidemiology and highlights recent major results in the field. While network-based models have desirable and complex properties that cannot be replicated by equation-based models, network models are computationally challenging and often result in high-dimensional data analysis (see Pellis et al. (2015)). For general study of the structure and dynamics of network models, as well as their applications in various disciplines, Boccaletti et al. (2006) and Newman (2003) are excellent reads.

3.1 Risk Network and Cyber Insurance

Böhme et al. (2017) provided a framework of cyber risk modeling and assessment. In contrast with the frequency and severity analysis of conventional risks, cyber risk is analyzed in the cybersecurity literature by a chain of causality from cyber risk factors such as threats, vulnerabilities, controls, and assets to financial losses. Each cyber incident is viewed as a threat acting on an objective's vulnerabilities. Companies can place technical controls to remove or reduce vulnerabilities. An incident is turned into a loss when attacks hit assets. The contagion of cyber risk is characterized by a network model, under which each firm is represented by a node and all nodes are interconnected by physical and social links. An example of cyber risk dependency is the propagation of malware through a network of interconnected computers. The specific dependence structure can be modeled by graph topology. Such topological models originated in epidemiological literature to model the transmission of infectious diseases and are extended in the context of cybersecurity to the spread of computer virus.

The causality approach based on network models, which is well studied in the cybersecurity literature, is fundamentally different from the trend analysis based on historical claims data in the actuarial literature. It has been argued in Böhme et al. (2017) that historical data is of limited use due to the fast-changing nature of cyber technology development. Therefore, further research is needed to combine actuarial analysis with cybersecurity framework in cyber risk quantification and modeling.

4. AGENT-BASED MODELS

The decisions of one person in a group can vastly affect the final outcome of the whole group. Therefore, when it comes to a community-scale activity, it is important to consider the competing strategies among all individuals. For those scenarios, an agent-based model (ABM) performs exceptionally well. Formally, it is a data-driven, simulation-based model that captures individual (or agent) level actions and view the community as a complex adaptive system. Zhou (2013) pointed out that ABMs bring insight of how to model agent's behavior, how to understand the learning process of an agent, and how to quantify the complex interactions among all the agents. Palin et al. (2008) elaborated on the main features of ABMs, namely (i) there are finitely many types of agents following different rules, (ii) agents are allowed to learn, combine, and evolve during the process, (iii) there will be positive or negative feedback from the agents' action, (iv) agents make contact with the environment and the environment will also be affected by the actions of the agents. As noted in Niu (2016), there are three types of ABMs, namely homogeneous ABM (where the agents are uniform and follow the same decision rules but have different parameters), heterogeneous ABM (where there are different types of agents communicating within the same environment), and agent-based queuing model (ABM that applies queuing theory).

4.1 Insurance Application

In order to make a comparison among insurance companies with similar business volume, Taylor (2008) built a dynamic model based on the model framework that each insurer is an agent. Moreover, applying advanced machine learning technique, Parodi (2012a, b) categorized the actuarial problems into different machine learning frameworks and applied computational intelligence methods to insurance business.

In light of the current pandemic, we propose a possible case study using ABMs. Suppose the insurer is interested in the number of confirmed and death cases in the United States for the coming fall. To model complex human behaviors and social phenomenon such as stay-at-home orders, social distancing, and mask wearing culture, ABMs are robust enough to incorporate these features. Historical data can be used to estimate the parameters in the ABM and subsequently one can simulate multiple scenarios based on these fitted parameters. Circling back to the insurance example in Section 2.1, these results can then be used to analyze how much reserve the insurance company will need for future benefit payments and what is the fair premium amount for the policyholder to pay corresponding to each level of the reproduction number.

5. CONCLUSIONS

This essay highlights the interplay between actuarial science and epidemiology and provides literature review of recent results. The marriage of modeling techniques from the two areas may revolutionize the way we quantify and model epidemic insurance and the way we organize cyber insurance. We hope to point to new sources of inspiration for actuaries working in areas of risk modeling and assessment of dynamically evolving populations. We envision actuaries adding these epidemiological models and techniques into their growing toolkits for risk modeling and analysis.

Acknowledgment

This research is supported by a grant by the Canadian Institute of Actuaries in response to COVID-19 research and by an endowment from the State Farm Companies Foundation.

6. REFERENCES

[1] Anderson, R.M. and May, R.M., *Infectious diseases of humans: dynamics and control*, Oxford university press, 1992.
[2] Andersson, H. and Britton, T., *Stochastic epidemic models and their statistical analysis*, Springer Science & Business Media, 2012, Vol. 151.
[3] Billard, L. and Dayananda, P.W.A, "A multi-stage compartmental model for HIV-infected individuals: I–Waiting time approach," *Mathematical Biosciences*, 2014, Vol. 249, 92-101.
[4] Billard, L. and Dayananda, P.W.A, "A multi-stage compartmental model for HIV-infected individuals: II–Application to insurance functions and health-care costs," *Mathematical Biosciences*, 2014, Vol. 249, 102-109.
[5] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.U., "Complex networks: Structure and dynamics," *Physics reports*, 2006, Vol. 424, No. 4-5, 175-308.
[6] Böhme, R., Laube, S., Riek, M., "A fundamental approach to cyber risk analysis," *Variance*, 2017, Vol. 12, No. 2, 161-185.
[7] Brauer, F., and Castillo-Chavez, C., "*Mathematical models in population biology and epidemiology*," Springer, 2012, Vol. 2.
[8] Brauer, F., "Mathematical epidemiology: Past, present, and future," *Infectious Disease Modelling*, 2017, Vol. 2, No. 2, 113-127.
[9] Chen, X., Chong, W.F., Feng, R. and Zhang, L., "Pandemic risk management: resources contingency planning and allocation," 2020.
[10] Diekmann, O., Heesterbeek, H., and Britton, T., "*Mathematical tools for understanding infectious disease dynamics,*" Princeton University Press, 2012, Vol. 7.
[11] Feng, R., and Garrido, J., "Actuarial applications of epidemiological models," *North American Actuarial Journal*, 2011,

Vol. 15, No. 1, 112-136.
[12] Fine, P., Eames, K., and Heymann, D.L., "“Herd immunity”: a rough guide," *Clinical infectious diseases*, 2011, Vol. 52, No. 7, 911-916.
[13] Kermack, W.O., and McKendrick, A.G., "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, 1927, Vol. 115, No. 772, 700-721.
[14] Lefèvre, C. and Picard, P., "Final outcomes and disease insurance for a controlled epidemic model," *Applied Stochastic Models in Business and Industry*, 2018, Vol. 34, No. 6, 803-815.
[15] Lefèvre, C. and Picard, P., "A general approach to the integral functionals of epidemic processes," *Journal of Applied Probability*, 2018, Vol. 55, No. 2, 593-609.
[16] Lefèvre, C. and Picard, P., and Simon, M., "Epidemic risk and insurance coverage," *Journal of Applied Probability*, 2017, Vol. 54, No. 1, 286-303.
[17] Lefèvre, C. and Simon, M., "Cross-infection in epidemics spread by carriers," *Stochastic Models*, 2018, Vol. 34, No. 2, 166-185.
[18] Lefèvre, C. and Simon, M., "SIR-Type Epidemic Models as Block-Structured Markov Processes," *Methodology and Computing in Applied Probability*, 2019, 1-21.
[19] Newman, M.E., "The structure and function of complex networks," *SIAM review*, 2003, Vol. 45, No. 2, 167-256.
[20] Niu, G., "Actuarial Application of Agent Based Modeling," 2016, Doctoral dissertation, University of Connecticut.
[21] Nkeki, C.I. and Ekhaguere, G.O.S., "Some actuarial mathematical models for insuring the susceptibles of a communicable disease," *International Journal of Financial Engineering*, 2020, Vol. 7, No. 2, 2050014.
[22] Palin, J., Silver, N., Slater, A., and Smith, A.D., "Complexity Economics Application and Relevance to Actuarial Work," In *Report of a working party of the UK actuarial profession. Paper presented to the Finance Investment and Risk Management Conference*, 2008.
[23] Parodi, P., "Computational intelligence with applications to general insurance: a review-I. The role of statistical learning," *Annals of Actuarial Science*, 2012, Vol. 6, No. 2, 307.
[24] Parodi, P., "Computational intelligence with applications to general insurance: a review-II. Dealing with uncertain knowledge," *Annals of Actuarial Science*, 2012, Vol. 6, No. 2, 344-380.
[25] Pastor-Satorras, R., Castellano, C., Van Mieghem, P., and Vespignani, A., "Epidemic processes in complex networks," *Reviews of modern physics*, 2015, Vol. 87, No. 3, 925.
[26] Pellis, L., Ball, F., Bansal, S., Eames, K., House, T., Isham, V., and Trapman, P., "Eight challenges for network epidemic models," *Epidemics*, 2015, Vol. 10, 58-62.
[27] Perera, S.S.N., "An Insurance Based Model to Estimate the Direct Cost of General Epidemic Outbreaks," *Int. J. Pure Appl. Math*, 2017, Vol. 117, No. 14, 183-189.
[28] Shemendyuk, A.A., Chernov, A.A., Kelbert, M.Y., "Fair Insurance Premium Level in Connected SIR Model under Epidemic Outbreak," *arXiv preprint. arXiv:1910.04839*, 2019.
[29] Taylor, G., "A simple model of insurance market dynamics," *North American Actuarial Journal*, 2008, Vol. 12, No. 3, 242-262.
[30] Zhao, S., Lin, Q., Ran, J., Musa, S.S., Yang, G., Wang, W., Lou, Y., Gao, D., Yang, L., He, D., Wang, M.H., "Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak," *International journal of infectious diseases*, 2020, Vol. 92, 214-217.
[31] Zhou, F., "*Application of Agent Based Modeling to Insurance Cycles*," 2013, Doctoral dissertation, City University London.

Auto Insurance: Strategic Shift Required for Acquiring and Retaining the Right Customers in a Post COVID-19 World

Swarnava Ghosh and Aditya

Motivation. This paper was written in response to a 'Call for Papers' on COVID-19
Method. This essay relies on external research and what we have experience in interacting with our insurance clients.
Conclusions. Devising the best pricing strategies for customers will be critical for success particularly in regard to auto insurance premiums, and insurers will have to apply sharply focused customer segmentation in order to develop relevant auto insurance marketing strategies that resonate on an individual level in the post COVID world.

Keywords. Pricing Models, Recovery, Macroeconomic Data, Segmentation

The impact of the COVID-19 outbreak has been devastating for the global economy, and with many countries in extended lockdown, there has been a seismic shift in customer behavior and business operations. As of July 10, US auto insurers such as Allstate, Farmers Insurance, Geico, The Hartford, Liberty Mutual, Progressive, Nationwide, State Farm, and USAA are expecting to return $14 billion to policyholders via refunds, discounts, dividends and credits.[1] Auto Insurance market is likely to see significant changes to the current business models in the coming months, with significant impact on revenues across the industry such as decrease in premium volume, delay in payments without penalty, and change in valuation and loss-recognition systems.

There are several factors at play contributing to a decline in revenues for the sector. The first is a sharp decline in automobile sales leading to a fall in number of new auto insurance policies. Vehicle sales in US for August'2020 declined close to 20% (YoY) whereas the total January-August 2020 sales were down by 23%. [2] Insurance companies oversee a dramatic change in customer behavior. With partial or full loss of income, people will be more hesitant to spend, leading to a sharp decline in automobile sales and a reduction in the number of new insurance contracts.

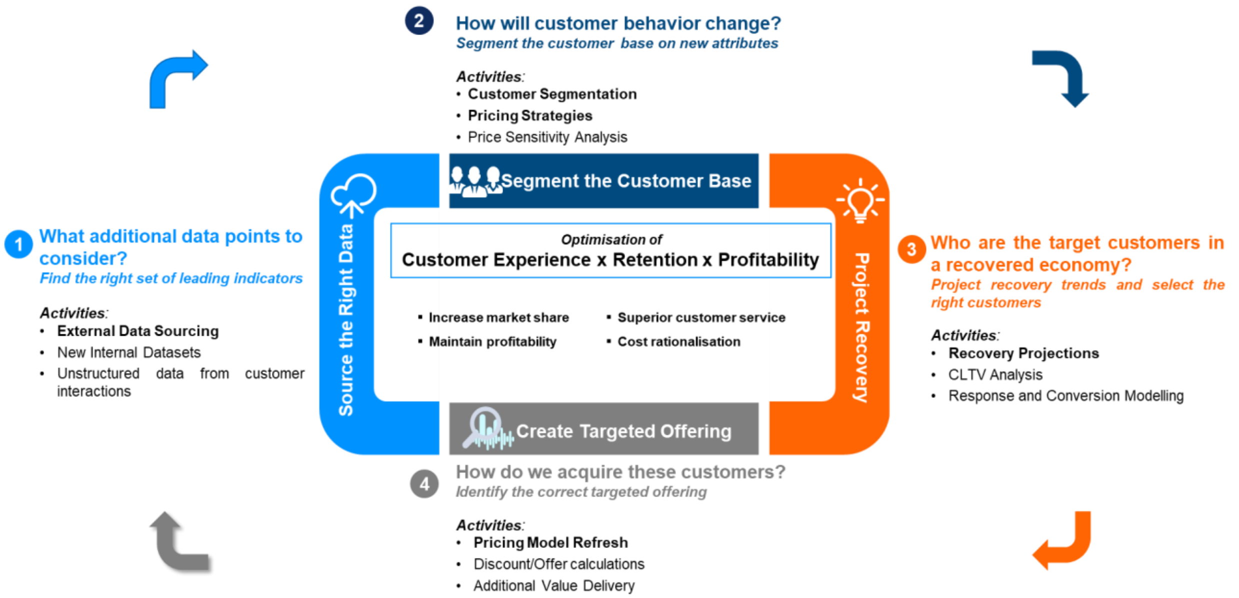
Secondly, with the world in lockdown and most offices adopting work from home policies, people are going out only for essentials. This has drastically reduced travel and automobile usage, For example, in the US, Cumulative Travel for 2020 changed by -14.5%.[3]

Thirdly, the mandatory grace period provided by insurance regulators for the payment of premium installments, can lead to more defaults in payment. This has also in turn reduced the amount of premiums received for auto Insurance. For instance, in the US, the Washington Office of Insurance Commissioner has provided grace periods for nonpayment of premiums and waived the fees associated with non-payment of premiums.[4]

To mitigate the impact of these revenue-reducing drivers, auto insurers need to make systematic changes in their current business strategy and adopt robust measures to optimize customer experience, minimize losses from new business, and improve retention while maintaining profitability. Some key objectives that these measures should address are:

- Defending and growing market share by retaining and acquiring customers
- Offering superior customer service
- Improving loss ratios
- Lowering cost of acquisition and servicing

Insurers can meet these objectives by following a four-step strategy driven by effective use of data and analytics across the value chain of the organization.



Step1: Sourcing the Right Data

It is critical to bring in the right data sources (both internal and external) so that insurers can execute the next 3 steps and make timely strategic decisions based on key recovery indicators. COVID-19-specific insights can be extracted using internal data points like telematics data and driving behavior, customer interactions, and Mid Term Adjustments during the pandemic, and by categorizing customers into sub-groups like essential/non-essential workers. External data sources such as mobility data, macroeconomic indicators, consumer behavior data, business activity, and recovery trends can be used to enrich the existing models for better business insights.

For instance, movement of people is highly correlated with vehicle usage, hence, mobility data can be a leading indicator of economic and business activity recovery. Some of the public sources which can be utilized for the same include Google Mobility Data, Apple Mobility Report etc.

Similarly, since forecasts for major macroeconomic indicators are widely tracked, insurers can use them to predict how Auto Insurance premiums may behave over the years. Employment rate has the highest correlation with Auto Insurance GWP, but it is a lagging indicator. Household Disposable Income mimics the GWP numbers very closely.

Likewise, based on how businesses perceive the economic threat of the COVID-19 crisis, there will be a direct impact on the volume of exports and imports, which will in turn impact the auto insurance business as well. Therefore, the Business Confidence Index can be leveraged here as it has the highest correlation with auto insurance GWP and its behavior over the years has closely resembled GWP trends.

Customer behavior and sentiment can be tracked using indexes like Consumer Confidence Index (CCI) and Consumer Price Index (CPI), which show the degree of optimism consumers feel surrounding the country's economy. Analyzing consumer behavior and sentiment can help insurers price accordingly and improve loss recognition models. In particular, CCI is a good indicator to feed into pricing models as it has a strong correlation with GWP numbers and closely resembles consumer's behavioral changes due to the pandemic. Other external indicators such as employment rate and household savings can also be used.

Step 2: Segmenting the Customer Base

In the post pandemic world, it is critical that insurers identify the right customers for their products to match the risk appetite and offer the best combination of channel, proposition, risk coverage, and customer experience. Insurers should be looking at their existing and prospective customer base and re-evaluating their target segments. It is essential to evaluate the needs and preferences of customers at each stage of their journey, and correlate these across marketing, sales & service, underwriting, policy administration and claims.

The fundamentals of segmentation for auto insurance customers have not changed, but several parameters have now become key to the segmentation process in a post COVID-19 world. These key parameters will be governed by requirements for travel and vehicle usage, relative risk of infection and financial strength of certain groups of customers.

With many people furloughed or out of jobs, employment status and type have become key categorical variables for segmentation. A distinction between essential versus non-essential workers has also become extremely critical with respect to future auto insurance requirements. As an example, customer segmentation using job category could look like:

a. Increased Mobility – Frontline Workers (24%)

b. Quick Return to Normal – Outdoor Activities (47%)

c. Slow Recovery – Hospitality & Travel (10%)

d. Work From Home – IT & Professional Services (19%)

A combined financial stress score created for the financial stress of individuals can be another key indicator for segmentation. It can consider different parameters such as affluence in the area of home address, primary source of income, job profile, stability of the employment sector, previous payment history (including any deferred payments during COVID19 crisis). This score can also help determine the propensity to buy add-on coverages and premium products.

Lastly, age is also a critical segmentation parameter. Since the impact of the virus is quite disproportionate based on different age groups, certain age groups (mostly above 50) are expected to stay in isolation longer and hence will have significantly reduced demand for auto insurance, which will also impact claim scenarios.

These new segmenting attributes could drive different marketing, servicing and pricing strategies for different groups of customers based on their incremental or reduced driving needs, propensity to buy different coverages and their unique risk factors. When using these attributes, one must be careful to avoid sensitive attributes in pricing models that might lead to discriminatory pricing. However, it is critical to align these segments to the right pricing strategy. A three-tier pricing strategy will align with most customer needs. This includes usage-based pricing, traditional pricing, and discounted pricing.

1. **Usage Based Pricing:** A low fixed premium with a variable usage-based premium calculated every cycle.

2. **Traditional Pricing:** The standard pricing strategy with a fixed premium for standard coverage allocated over the policy period.

3. **Discounted Pricing:** A fixed premium pricing with a lower premium providing only basic coverage.

Once a new customer segment is created, it is also important to perform price sensitivity analysis for each segment to position different products and coverages appropriately surrounding that sensitivity.

Step 3: Projecting Recovery

In the past, auto insurance premiums have shown strong correlation with a few macroeconomic indicators, business activities and customer indicators. Today, insurers must track these indicators in real time to project what the future holds in their road to recovery.

The data can be used for projecting premiums, recovery trends, and market sentiments. For example, it has been seen that employment rates have a direct correlation to household spending and an indirect correlation to household savings - which further translates to auto sales and impacts mew auto insurance contracts.

Additionally, due to changes in risk factors and the financial appetite, there will be a significant shift in the lifetime value of different customer segments. This will require an overlay of different behavioral and financial trends over existing CLTV models to make them more robust and dynamic so that they are able to capture COVID-19 related uncertainties.

To increase CLTV, there will also be a shift towards a bundle approach rather than singular products. The bundle approach will streamline the process for customers as they can buy different policies from a single insurer.

Furthermore, customers from different segments will be impacted differently by COVID-19, and their risk appetites and financial outlooks will change. This shift will have a major impact on response and conversion models that have been built on historical data and which are becoming obsolete.

Insurers will need to assess the impact of the new normal and incorporate new variables and data modeling practices. This will see inclusion of real time data, application of telematics and use of advanced Machine Learning techniques like Generative Adversarial Networks (GANs).

Step 4: Creating a Targeted Offering

In the long term, the behavioral and operational changes associated with the pandemic will impact not only the way fundamental risk assessment is done for technical pricing, but also the way retail price and premiums are determined by insurers. The size of the auto insurance market will likely contract, and insurers will be looking to defend and grow their market share. The industry will need to constantly innovate and come up with novel targeted offerings that can help acquire and retain the most profitable set of customers.

The starting point of this will be a refreshment of pricing models. The frequency of claims is trending downwards in recent months, but the claims are more severe in nature; [5] which means that many loss cost models will require refreshing in order to incorporate the new scenarios. Therefore, personal line insurers will need to incorporate newer data points emerging from greater use of technology, including telematics and IOT devices.

Additionally, while Machine Learning and AI have made significant progress in areas such as claims and marketing, adoption has been quite low in pricing due to challenges around definition and reinforcement of biases in machine learning models. Insurers are now able to capture significant amount of customer data which can be combined with advanced machine learning models to better evaluate customer risk profiles and provide customized prices to their customers. The current traditional GLM models can go only so far due to limited accuracy and prediction power. .This will help insurers reduce losses and improve customer retention.

Pricing teams on the other hand will need to adopt a rapid experimentation approach for getting results faster, while also experimenting with different methodologies in an agile manner.

In terms of refresh frequency, typically core pricing models are not refreshed frequently and only marginal changes are done. However, in a rapidly changing insurance landscape, it is imperative to increase the frequency of model refreshes.

Also, for machine learning based pricing model deployment, it is critical to set up a flexible and scalable technology architecture that can ingest data and provide outputs in near real time in order to satisfy the complex requirements of personal lines insurers.

The second major point is the new insurance offers/discounts that will be provided to the customer. This is vital because how insurance firms treat their customers now is likely to create a long-lasting perception in the market. Insurers that understand the plight of the people and implement fair and balanced reforms in their pricing and premium practices will be rewarded in the long run. However, there are multiple methods of calculating and offering the discounts to the end customer. Some of the common methodologies are:

- Flat Discounts
- Flat Personalized Discounts
- Premium Proportionate Discounts
- Premium & Claim Proportionate Discounts

Lastly, Insurers can take other proactive steps to create additional value for their customers, such as utilizing any spare capacity in their sales, service and claim contact centers to keep in touch with tenured and elderly customers in these difficult times to inquire about their well-being. Making the necessary changes in customer journeys to make them more streamlined and digital is also an additional value add.

In conclusion, COVID-19 is unlike any challenge humanity has faced in the recent past, and the economic impact has been unprecedented. Like all sectors, insurance recognizes the need to be agile and to use new data to project recovery so that it can understand the behavioral and financial changes customers are experiencing. Devising the best pricing strategies for customers will be critical for success particularly in regard to auto insurance premiums, and insurers will have to apply sharply focused customer segmentation in order to develop relevant auto insurance marketing strategies that resonate on an individual level.

REFERENCES

[1] Kiger, Patrick, and Andy Markowitz, "Big Auto Insurers Phase out Refunds for Policyholders,"
[2] Madhok, Aman, "Weekly Update: COVID-19 Impact on Global Automotive Industry,"
[3] Traffic Volume Trends September 2020 from Policy of Governmental Affairs Office of Highway Policy Information
[4] Jaconetti, Annie, "Washington State Insurance Commissioner Orders Insurance Premium Grace Period,"
[5] Masri, Lena, "Car crashes deadlier as drivers speed during lockdowns,"

Abbreviations and notations

CLTV, customer lifetime value

GLM, generalized linear models

IoT, internet of things

AI, artificial intelligence

Biography(ies) of the Author(s)

First Author is Swarnava Ghosh (Senior Engagement Manager – Analytics): A dedicated analytics professional with 10+ years of experience in global engagement management, analytics project execution, business case development, building data and analytics strategy and risk management in insurance. He has led teams on numerous engagements involving interaction with senior stakeholders in Pricing & Underwriting, Claims, Finance, Risk, Sales & Marketing, Operations and other business functions. He has extensive experience of setting up and running Analytics Centre of Excellence at scale across UK, US, India and Philippines.

Second Author is Aditya (Assistant Project Manager – Analytics): A dedicated analytics professional with extensive knowledge of analytics model development and implementation in BI platforms for decision support. He has worked on managing multiple stakeholders, development and deployment of analytical models, implementation of efficient solutions and driving data science enabled growth to solve challenging business problems for key major P&C insurers in domains like Marketing, Claims and Underwriting.

Everyone Can Assist in the Battle Against COVID-19: A CAS Member's Experience

Kwok Ching Ng

The author of this short essay is a member of the State of Vermont COVID-19 modeling team. The main purpose of this essay is to share the author's experience serving in the team and to show how, with some initial investments in learning the basic domain knowledge, any actuary, or person with quantitative and/or critical thinking skills, can add value in the fight against COVID-19 or other pandemics that might arise in the future. The essay also describes the composition of the Vermont modeling team and what it does in assisting the Governor and his cabinet in making COVID-19 policy decisions. Opinions expressed in this essay are strictly the author's.

THE VERMONT COVID-19 MODELING TEAM

As of mid-October 2020, the State of Vermont arguably had the best track record amongst all 50 states since the beginning of the pandemic for the control and mitigation of COVID-19. It had the lowest COVID-19 death rate, the lowest infection rate, and some of the lowest positivity rates amongst all 50 states. "This should be the model for the country, how you've done it," Dr. Anthony Fauci said in a recent video press conference that Vermont Governor Phil Scott held with the state's media. Dr. Fauci also said that being a rural state does not automatically guarantee a better outcome (1) (2). In fact, we could all see how COVID-19 proliferated in the Upper Midwest and other rural states in September and October. The credit for Vermont's success can be attributed to the Governor and his staff, which early on decided to follow science and advice from State Health Commissioner Dr. Mark Levine, State Epidemiologist Dr. Patsy Kelso, and others including the faculty of the University of Vermont Medical School. The Governor held press conferences at least two times a week in which he and the heads of various state government agencies/departments informed the public of what had been happening and the reasons for imposing specific interventions, as well as reasons for subsequent relaxations. Most of all, Vermonters overwhelmingly trust their state government to make the right decisions and largely comply accordingly.

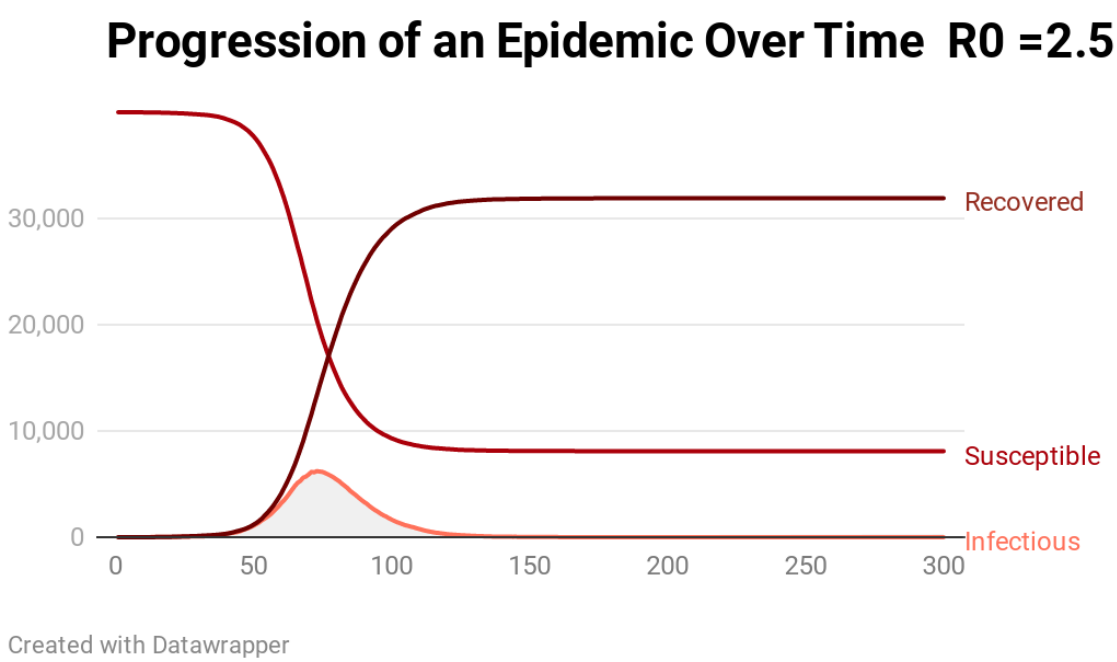
Because all the state's epidemiologists were more than fully occupied with their duties and could not spare time or resources to do COVID-19 modeling, the Commissioner of Financial Regulation, Mike Pieciak, a securities lawyer by training, was asked by the Governor to head up the COVID-19 modeling team in March. A few weeks later, the Deputy Commissioner of Insurance, Kevin Gaffney, who hired me six years before, recommended to Mike that I join his team (while retaining my existing duties). At that time, the only other member of the team was Isaac Dayno, who is a Harvard graduate trained in the liberal arts but with experience in organizing facilities for people with HIV or hepatitis, as well as experience in running political campaigns. As I soon learned, his skillset is important because messaging is a very important part of any public health campaign. A few weeks later, we hired a rising senior from Yale, Ryan Taggard, majoring in Mathematics and Data Science. Ryan's Python programming skills, especially in data collection and graphics, quickly proved to be indispensable in making Mike's weekly presentation to the public come "alive" – with supporting data from around the country. A few weeks after Ryan's arrival, an epidemiologist, Mary-Kate Mohlman, finally joined our team on an as-available basis. Needless to say, having her assistance was a big win. Later still, we began to work with John Adams, the Director of Vermont Center for Geographic Information, to provide us with periodic mobility graphical analyses using SafeGraph mobile phone data. The resolution of SafeGraph's data is an order of magnitude higher than those in Google Mobility Reports.

The roles and tasks of the modeling team evolved over time. From the start, Commissioner Pieciak sought advice from various reputable COVID-19 modeling teams around the country – including Columbia, Northeastern, IHME and Oliver Wyman, among others. All those modeling teams provide state by state projections of COVID-19 confirmed case counts and death counts with intervals of a few weeks to a few months. We would feature one or two modelers' projections in the weekly press conference presentation, along with COVID-19 related Vermont data. Of particular importance were COVID-19 projections under various degrees of non-pharmacological intervention scenarios – such as closure of schools, child-care programs, restaurants, bars, gyms, salons and spas; compliance with Stay Home Stay Safe; and suspension of all in-person business operation, etc., for all businesses and not-for-profit entities.

THE SIR MODEL

Right from the start, I knew I had to acquire as much domain knowledge of epidemiology as quickly as possible. I began to read about SIR models – from basic introductions to dozens and dozens of journal papers in epidemiology. It soon became clear to me that all the COVID-19 non-pharmacological interventions that the country, and in fact the whole world, were considering using can be traced back to, or get their hints from, the SIR models.

I learned that there are many infectious disease models; some are deterministic (SIR, SEIR), and some are stochastic. The most complex ones are network models. The most common one, the SIR model, has been around since the 1930s but still works amazingly well. Common graphical illustration of an epidemic from the beginning to the end based on a simple SIR framework looks like the following:



The model is a dynamic system governed by the following differential equations (5):

$$\begin{cases} \dfrac{dS}{dt} = -\beta \dfrac{S(t) \times I(t)}{N} \\ \dfrac{dI}{dt} = \beta \dfrac{S(t) \times I(t)}{N} - \gamma I(t) \\ \dfrac{dR}{dt} = \gamma I(t) \end{cases}$$

where N is the population size of a closed population, β is the effective contact rate of the disease, γ is the decay rate of the disease, and S(t), I(t), and R(t), respectively, represent the number of susceptible, infectious, and removed (includes those recovered or dead) individuals in the population at time t.

The effective contact rate β is determined by $\beta = \mathrm{T} \times \mu$, where T is the number of exposures per unit of time, and μ is the probability of infection from each occasion of exposure. Clearly, non-pharmaceutical interventions such as social distancing, mask wearing, frequent hand washing, closure of businesses, closure of schools, quarantine, etc., are all efforts to either lower T, or μ, or both, thereby reducing β, which in turn lowers the rate of infection or disease transmission – according to the equation for dI/dt above.

I have included some additional information about the SIR model in the Appendix of this essay. There are many journal articles on infectious disease modeling. For example, see (5) for the basic SIR model, (6) for an application of an SEIR model, (7) for an example of a Markovian stochastic SIR model, and (8) for a lengthy discussion of SIR Network models (not an easy reading, though).

A NEW METHOD FOR ESTIMATING ACTIVE (INFECTIOUS) CASE COUNT

At the early stage of the pandemic, measures of an epidemic's reproduction number (commonly referred to as R_0 and R_t) were of great interest to many people. I read papers on the subject and then requested the state's COVID-19 Data Team (who are all epidemiologists) in the Vermont Health Department to provide me with "infector-infectee pairs" data. This information would allow me to create a "Serial Interval Distribution" – one of the input requirements for arriving at estimates of R_t. Basically, the data gives a collection of the number of days between symptom onset of a given infector and symptom onset of the person infected by that infector. That is, the data is a (surrogate) frequency distribution of the number of days required for an infector to pass the virus to an infectee - usually in a family setting or in the workplace. I was provided with 50 pairs (many journal articles use 25 to 40 pairs) and found that a two-parameter Gamma distribution provides a good fit to the data. It should be noted that epidemiologists tend to use Gamma or Lognormal for Serial Distributions. Weeks later I proposed to "off-label" use this distribution as the primary tool for measuring infectiousness by state and counties within a state, turning it into a critical tool for Vermont's travel quarantine policy and for assessing how infectious Vermont's neighboring states are at any point in time.

During the middle of May, when the Northeast states' infection rate began to improve, Vermont Governor Scott wanted to partially open Vermont to tourists from the Northeast. He and his cabinet, with consultation from the Health Department and Commissioner Pieciak, settled on a threshold of 400 "active cases" per million population. Residents of counties with active cases lower than that could travel to Vermont without quarantine. It was then up to the modeling team (us) to figure out how to calculate active case counts. This was not a trivial task. At the time, many states did not publish active cases by county. Even more concerning, we did not know how each state defines/computes active cases, as there was and still is no uniform standard. What we needed was a uniform measuring methodology to be applied to all states and all counties within states. Our own Health Department's procedure is to consider an infected person "active" until 30 days after symptom onset, or sooner if the infected person was clinically determined to have "recovered." In their modeling, Oliver Wyman, on the other hand, simply considers a person "active" during the 14-day period from onset. We were about to adopt

Oliver Wyman's method, but the "Serial Interval Distribution" clearly tells us that an infected person is most infectious during the early days of symptom onset, becoming progressively less so. By the 10th day after onset, the person's infectiousness becomes very low. Counting a person infected yesterday and someone who was infected 10 days ago both as "active," is not a good measuring system for the purpose of differentiating degrees of infectiousness among counties. On the other hand, using the area under the Serial Interval Distribution as "a unit of life-time infectiousness" (or life-time viral load for that person's COVID-19 epoch), a newly infected person would be carrying the highest amount of virus to be shed – 1.00 unit. A person infected 5 days ago would have significantly less "life-time viral load" remaining to be shed, say 0.4 unit. Hence one minus the cumulative distribution (1 - F(t)) of the Serial Interval Distribution (the fitted Gamma distribution) corresponding to the number of days (t) after a person's symptom onset, gives us the life-time viral load unit remaining for that person's COVID-19 epoch. As it turns out, this approach for measuring "active cases" had never been used anywhere before.

Initially, this approach was met with some skepticism, as it should. However, people gradually began to see the merits of the approach. Later, we presented it to two professors at Columbia and Northeastern, respectively, and got concurrences from them. We also got very positive comments from public health modelers in Massachusetts. I also tested the methodology using an SIR simulation model in the public domain (SimInf), and found the two produced quite similar, but not exact, results in terms of number of infectious people remaining at successive points along the timeline. As I read more journal articles, I found that an even more theoretically correct distribution for our purpose would be the convolution of the Serial Interval Distribution and the incubation period distribution. But it has been said, perfect is the enemy of good. One week after I submitted the initial version of this essay to the CAS, we received an unexpected e-mail from Professor Ronald Lasky at Dartmouth College informing us that Dartmouth would like to use our approach for estimating active cases. It was gratifying to see our methodology being adopted by an institution such as Dartmouth. See links in the References section to the documentation of our methodology (3) and our weekly presentation (4).

RANDOM TESTING NOT EFFECTIVE IN IDENTIFYING THE INFECTED BUT ASYMPTOMATIC

Another "discovery" I made during the last several months was, while random COVID-19 surveillance testing could tell us the prevalence of COVID-19 infections in a geographic area, it has a very low expected positivity rate and therefore could be diverting limited resources away from the very important tasks of identifying infected individuals who are symptomatic for quarantine, and from the state to follow up with contact tracing. On the other hand, the positivity rate from testing symptomatic people for COVID-19 is almost always an order of magnitude higher than from performing surveillance testing on people with no symptoms, including asymptomatic individuals. This accidental "discovery" occurred when I was trying to help answer a critic of our active case estimation methodology. Using binomial approximation of Poisson, I concluded that:

With a relatively low population infection rate, conducting random surveillance COVID-19 testing has very little chance of identifying infected individuals, symptomatic or not. In the case of Vermont, the 7,000 tests performed per week in July were expected to only identify approximately two infected individuals or less, per week. Clearly, that would not be an efficient use of COVID-19 testing resources if test kits are in short supply, or future supplies are highly uncertain.

The main reason why Vermont had 30+ to 50+ positive cases identified per week in June to July was due to "self-selecting" or "self-selection" bias either because symptomatic individuals went to get tested/see a doctor, or contact tracing triggered the need for a test.

This highlights the reality that while surveillance testing is useful for determining the prevalence of COVID-19 in a state or a county, it offers very little help in identifying asymptomatic infected individuals for the purpose of stopping them from infecting others. This would still be true even if we increased total testing per week, say fivefold to 35,000, in Vermont. We will have to rely on other means to control spread by such individuals, such as requiring everyone to wear face masks because, by definition, we cannot tell that a person is infected if the person happens to be of the asymptomatic type. Only testing can tell that. However, we do not have the resources to test everyone who is asymptomatic on a regular basis. The CDC's best guess is that 20% to 50% of all those who are infected are asymptomatic.

I passed on my observations to the team/the Commissioner. A few days later July 24th, the Governor announced a mask wearing mandate - to be effective August 1st. It was something he had been reluctant to do for some time because he preferred educating the public over mandates. Also, during that same week, the Health Department stopped encouraging people with no symptoms to get a COVID-19 test, presumably to conserve test kits. I have no way of knowing the extent, if any, to which my conclusions contributed to those decisions. No doubt, a lot of considerations and inputs from experts and advisers were examined before such key decisions were made, including the fact that at that time COVID-19 was surging along the Eastern Seaboard, marching toward Vermont. The important thing to me was that those were right decisions.

OLIVER WYMAN AS VERMONT COVID-19 CONSULTANT

Even though Vermont experienced only small increases in confirmed COVID-19 cases after the Memorial Day and July 4th holidays, there were serious concerns about schools and colleges reopening as September was approaching. State officials saw elevated risks further down the road from large gatherings during major holidays such as Thanksgiving, Christmas, and New Year's; from visitors during the Vermont ski season; and from winter weather in general, when people stay indoors much more. By the middle of August, Commissioner Pieciak decided to enter into a formal agreement with Oliver Wyman (OW) to provide our team with COVID-19 analyses specific to Vermont in weekly video conference calls. This hugely increased our modeling bandwidth as by that time OW had already been providing COVID-19 consulting services to other entities across the country for several months. Their team consists of a well-known leader in financial industry modeling, an epidemiologist, a medical doctor, consultants from their Health & Life Sciences and Financial Services arms, plus software-programming staff. I found OW's analyses and advice helpful and generally well-thought-out, and innovative at times. I especially like their COVID-19 health risk scorecard by state and their analysis on the necessary conditions and timeline for the country to get back to normalcy. See a link to OW's COVID-19 projections by state (open access) in the References section (9).

In September, like many modelers, OW was developing tools for sizing up the potential impacts from school and college reopenings. OW provided some guidance and relevant information to us, but no projections. As time passed, it was becoming clear to all that, thanks to the efforts of all Vermont school districts, colleges and universities, and support from the state government, Vermont's K-12 and higher education were doing very well in absolute terms and, in comparison to Vermont's nearby states, in terms of having low positive COVID-19 case count per capita. The challenging thing for our team was in securing timely information from all the different school districts, colleges, and universities each week, which we then consolidated and presented at the Governor's press conference every Tuesday. Our focus next turned to Thanksgiving.

MODELING POTENTIAL IMPACT OF THANKSGIVING GATHERINGS

Around mid-November, the Governor's office asked Commissioner Pieciak to review a "COVID-19 Event Risk Assessment Planning Tool" that the Georgia Institute of Technology (Georgia Tech) made available online (10) (11). The Commissioner asked Ryan and me to review the tool and share our opinions with him. In addition, he also wanted to see if the tool could be helpful in analyzing scenarios around Thanksgiving.

My conclusion was that the risk assessment tool relies very heavily on one estimator: The probability that one or more attendees in a large gathering of size N are already infected with COVID-19 when they arrive. That probability is a function of gathering size and the state's COVID-19 prevalence on any given day. However, once the event size reaches 100% for the said probability for a given state, the tool provides no distinctions between all larger event sizes. In particular, it does not provide any framework for estimating how many new infections could take place during the gathering for progressively larger events, which is a very important consideration for risk assessment and planning purposes.

Conceptually, the missing piece could be approximately modelled by this formula: Expected Number of New Infections During the Gathering = N x P_{SCt} x ARN, where N is the gathering size; P_{SCt} is the probability of an attendee arriving on date t from state S, county C at the event, already infected; and ARN is the "Attack Rate,"- which represents the proportion of people expected to be newly infected during an event of size N. The next step was to estimate P_{SCt}. I chose to estimate that by the number of active (infectious) case count divided by population corresponding to state S, county C, on date t. Active case count came from the methodology described earlier in this essay. Estimating Attack Rates was much harder. Given that Thanksgiving was only seven days away, I decided to rely on empirical data as a starting point and then used an exponential curve in guiding my selections of AR for various sizes N. The empirical data came from various COVID-19 outbreaks traced back to large gatherings such as weddings, church services, birthday parties, etc., reported in the news over the previous few months. The size of the gathering and the number of attendees who got infected were included in those news reports.

For estimating the potential impact of large Thanksgiving gatherings on COVID-19 transmission, Commissioner Pieciak sent me historical surveys by Pew Research Center and YouGov of Thanksgiving family gathering sizes, as well as some less detailed surveys done earlier in 2020. The percentage of such gatherings with guests from out of state was available too. Some of those surveys were done at the national level, and some just for the Northeast Region. Such was the data we could obtain.

After merging and distilling the above information, I estimated the expected new COVID-19 infections due to Thanksgiving gatherings in 2020 - with and without invited guests, based on different assumptions about the extent of voluntary curtailment in inviting non-family members for Thanksgiving. The conclusion: new infections could be 4 to 6 times higher, relative to having no invited guests for Thanksgiving.

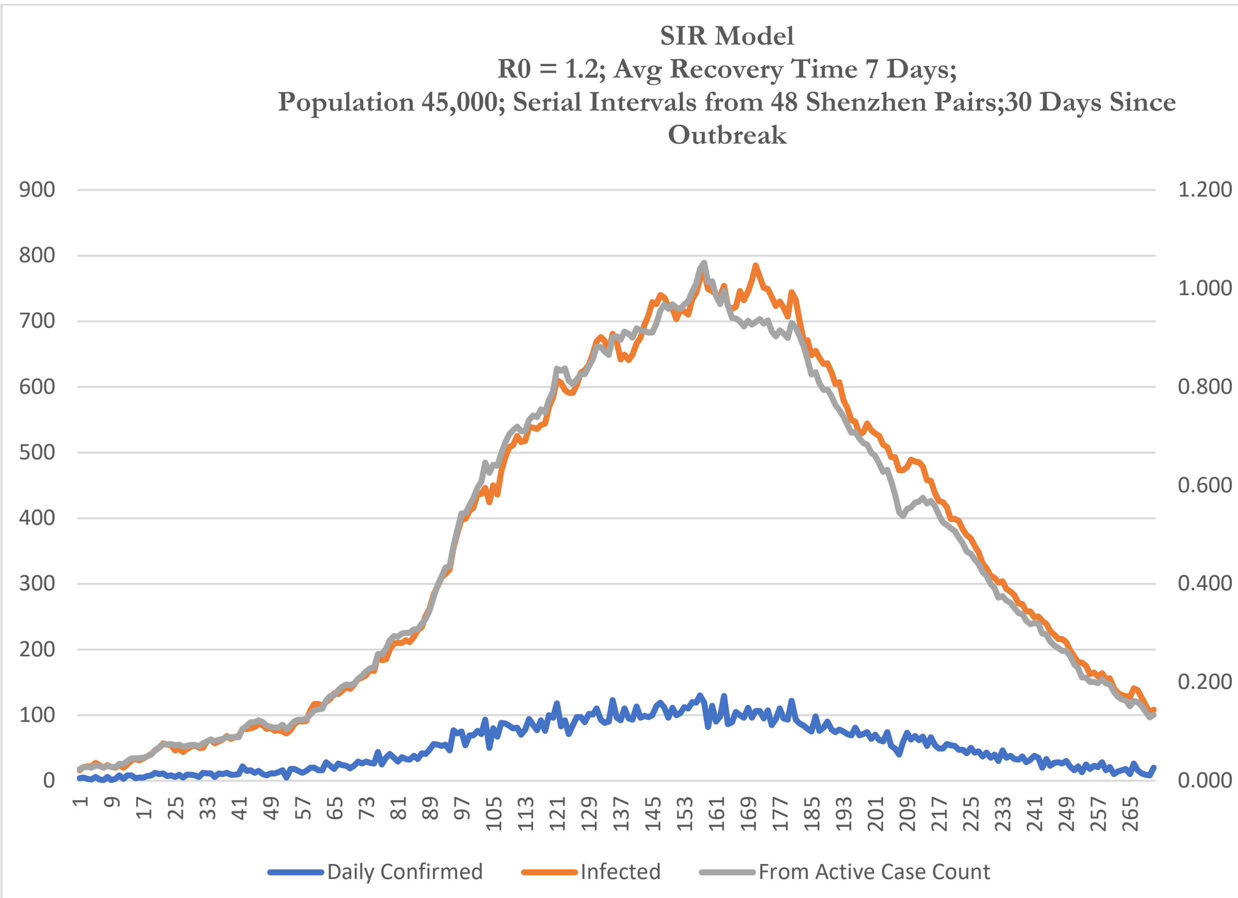
The Commissioner then asked me to send the model and results to Oliver Wyman for feedback. Two days later OW arrived with their own estimates, which turned out to be very similar to ours. Their approach was similar to ours too, except they used tools that they had already compiled/built for estimating the "Attack Rates" more scientifically: room dimensions, ventilation rate, duration of gathering, mask wearing, breathing volume rate, exhalation rate, etc., are in the model. See the link for "COVID-19 Indoor Safety Guideline" by researchers from MIT on how these considerations affect COVID-19 transmissions in the References section (12). Our findings were included in Commissioner Pieciak's presentation at the Governor's November 24th press conference. The intent was to reinforce the message, numerically, that the consequences from large Thanksgiving gatherings could be very severe, with the hope that most Vermont households would keep their gatherings small.

EVERYONE CAN ASSIST IN THE BATTLE AGAINST COVID-19

The modelling I described above could be comfortably performed by most actuaries. An actuary with more advanced skills than me in programming and modeling might have done more. Within our team, neither the Commissioner (a lawyer), nor Isaac, who has a degree in History, have much prior quantitative training. Yet they have a good grasp of the pandemic and are deeply involved with the state's daily battle with COVID-19. Clearly actuaries, statisticians, data scientists, economists, physicists, engineers … all have the capacities to help solve problems on the quantitative side of COVID-19. The world just does not have enough epidemiologists and doctors right now, so do not hesitate to take a deep dive into learning basic epidemiology and then help in any way that you can. We do not know what the future holds.

APPENDIX

Regarding the SIR model, I would suggest that once you have acquired some understanding of how the model works, say from (5), you would want to move on to read (or skim) as many other COVID-19 related articles on epidemiology as your schedule permits. Most epidemiology journals around the world offer free access to everyone during this pandemic. I would also highly recommend trying out the R package SimInf: An R Package for Data-Driven Stochastic Disease Spread Simulations. Vary the input SIR parameters and see how each epidemic unfolds. I used SimInf to test our Active (Infectious) Case Count methodology by running many hypothetical epidemics. The Infectious counts generated by SimInf and our Active (Infectious) Case Counts were in good agreement over the course of most hypothetical epidemics. The following is an example:



REPRODUCTION NUMBER R_0 AND R_t

The most comprehensive paper on Reproduction Number R is in (13) by The Royal Society. You would be better off spending time reading that paper than reading most other papers on R. I quote from the first page of the paper:

"R_0 represents the basic reproduction number, which is the number of secondary infections generated from an initial case at the beginning of an epidemic, in an entirely susceptible population. In contrast, R_t is the reproduction number at time t since the start of the epidemic. As more individuals are infected or immunised, R_t captures the number of secondary infections generated from a population consisting of both naïve/susceptible and exposed/immune individuals and therefore it both changes in value over time and will always be less than R_0."

It is important to know the same virus can have different R_0 depending on where the virus is taking hold. The R_0 for COVID-19 is expected to be different in the metropolitan areas of the US than, say, Fairbanks, Alaska, or Mongolia. This is because R_0 is partly determined by the biology of the virus, and partly driven by how people live and interact.

To estimate R_t, read the paper by Anne Cori et al. (2013) (14). There is also a corresponding online app for estimating R_t called EpiEstim App, and an R package called EpiEstim.

REFERENCES

(1) Dr. Facui spoke in the Vermont Governor's press conference on 9/15/2020: https://www.youtube.com/watch?v=axG1eJfXn00&list=PL-xsDpLCa0iRMj4dQrqum7uoXaQmqeGvT&index=11 (Dr. Fauci spoke 5 minutes into the conference)

(2) Which states had the best pandemic response, https://www.politico.com/news/2020/10/14/best-state-responses-to-pandemic-429376

(3) Vermont Modeling Team Active Case Count methodology: https://dfr.vermont.gov/document/travel-map-methodology

(4) Vermont Modeling Team presentation slides: https://dfr.vermont.gov/about-us/covid-19/modeling

(5) Ulrich Nguemdjo et al. (2020), Simulating the progression of the COVID-19 disease in Cameroon using SIR models: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0237832

(6) Alberto Gadio et al. (2020), SEIR Modeling of the Italian Epidemic of SARS-CoV-2 Using Computational Swarm Intelligence: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7277829/

(7) Henry Tuckwell & Ruth Williams (2006), Some properties of a simple stochastic epidemic model of SIR type: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7094774/

(8) Leon Danon et al. (2011), Networks and the Epidemiology of Infectious Disease: https://www.hindawi.com/journals/ipid/2011/284909/

(9) Oliver Wyman: https://pandemicnavigator.oliverwyman.com/forecast?mode=states&region=US_US-NY&panel=baseline

(10) Georgia Tech paper on COVID-19 events risk assessment: https://www.nature.com/articles/s41562-020-01000-9

(11) Georgia Tech online app for COVID-19 events risk assessment: https://covid19risk.biosci.gatech.edu/

(12) COVID-19 Indoor Safety Guideline: https://indoor-covid-safety.herokuapp.com/

(13) The Royal Society on Reproduction Number R: https://royalsociety.org/-/media/policy/projects/set-c/set-covid-19-R-estimates.pdf?la=en-GB&hash=FDFFC11968E5D247D8FF641930680BD6

(14) Anne Cori, et al. (2013), A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics: https://academic.oup.com/aje/article/178/9/1505/89262

ACKNOWLEDGEMENT

The author would like to thank the reviewers of this essay, Sandra Callanan, FCAS, and Denise Ambrogio, FCAS, for their meticulous edits, insightful comments and suggestions.

Examining the COVID-19 Impact on Human Capital

Mark Maenche, ACAS, MAAA, CIC, CRM

Abstract. The following essay is a response to the CAS call for essays on the topic of COVID-19. The essay focuses on the impacts that COVID-19 has on the broad categories of individual well-being and work. After illuminating these impacts, three areas of human capital management are suggested for the consideration of property-casualty leadership.

Keywords. COVID-19, Human Capital

1. INTRODUCTION

2020! Wow - what a year this is shaping up to be! Massive brush fires in Australia, the tragic death of an NBA superstar, a Presidential election…the list could go on and on. Looming over all the events that could be listed is the COVID-19 pandemic. Societal shutdowns and governmental stay-at-home orders have significantly affected the way businesses operate, including the Property and Casualty (P&C) insurance industry. P&C insurers are not only faced with protecting their financial statements but also managing the physical and mental well-being of their employees. The ramifications of COVID-19 on the P&C insurance industry's human capital will be felt for years to come and identifying solutions at this moment in time is a significant challenge for industry leaders.

The companies that comprise the P&C insurance industry employ approximately 650,000 people, according to the Bureau of Labor Statistics as of June 2020. These employees create economic value based on their experience and skill. This is the essence of human capital. It is an intangible asset made up of many different characteristics and qualities of the individual employees. Education, training, intelligence, health, and loyalty are all examples of such characteristics that may be contemplated when attempting to define human capital for a firm [1].

2. COVID-19 IMPACTS ON HUMAN CAPITAL

COVID-19 has caused the P&C insurance industry to encounter a diverse set of challenges that affect human capital. Human capital is intangible and difficult to quantify under ideal circumstances, making these new challenges presented by COVID-19 hard to assess and daunting to analyze. Consider these impacts on human capital:

2.1 Well-Being

In April and May of this year the American Psychological Association conducted a survey on stress which indicated that U.S. adults were approximately 20% more stressed this year as compared to the average stress level reported in 2019 [2]. Fear of themselves or a family member contracting COVID-19 and the government's response to the pandemic were significant drivers of this stress. Different employees will handle stress in a variety of ways, but it seems likely that the mental health effects could hinder productivity at work. An article in Human Resource Executive indicates that 62% of workers report losing at least one hour a day in productivity due to COVID- 19-related stress, with 32% losing more than two hours per day [3].

Virtually anyone whose job could be done from home is now working from home. The potential for employees to experience ergonomic problems has substantially increased. It seems reasonable to assume that most P&C insurance industry employers had not previously provided an optimal workspace setup for each employee's home. To immediately do so in light of COVID-19 is, most likely, cost prohibitive. Therefore, employees have been left to figure out what works best for them. How many people have been working from the couch or at the dining room table for the past six months? What is the number of Zoom meetings undertaken in less than ideal lighting?

A reduction in physical activity due to lockdowns and social distancing has undetermined long-term ramifications. A study published in June showed that the United States had a 15% decrease in step count within 15 days of the declaration of the pandemic [4]. The CDC recommends that adults get at least 150 minutes of moderate-intensity physical activity per week. This leads to many benefits as described by Janet Fulton, Ph.D., of CDC's Division of Nutrition, Physical Activity, and Obesity. She said, "Being physically active has many benefits, including reducing a person's risk of obesity, heart disease, type 2 diabetes, and some cancers. And on a daily basis, it can help people feel better and sleep better [5]." Decreased physical activity for employees can lead to potential increases in healthcare costs for employers. Sample text for this section.

2.2 Work

The switch to remove work opens the door to the potential for long-term working from home. This is a result of a number of factors: (1) Cost of real estate - since COVID-19 has forced so many within the P&C insurance industry to work from home, the need for physical real estate must be contemplated. (2) Social distancing requirements and limits on number of people inside buildings - gathering many employees at a single location is not allowed in some parts of the country and discouraged most everywhere else at this time. (3) Expansion of the talent pool – an insurer could theoretically hire from anywhere in the world. (4) Demand – current or potential employees may be unwilling to join the team unless remote work is an option. In order to retain/secure the best talent employers may need to continue using the work-from-home model. In May, performance marketing company Fluent published a study from data gathered online that indicated 59% of employees would continue to work from home once restrictions are lifted if given the option to do so [6].

The restrictions for social distancing have reduced the ability for P&C insurance industry professionals to get in front of customers due to limitations on travel. This includes salespeople meeting with buyers purchasing insurance policies, claims adjustors administering claims, and marketing representatives conversing with their clients. The decrease in sales contacts could cause repercussions as the sales funnel is populated with fewer buyers. Claim settlement times may increase affecting the timing of loss payouts. Thankfully, the internet has allowed industry professionals to continue to perform these functions, albeit in a limited fashion. With this transition P&C practitioners are having to learn new technological skills on the fly.

3. AREAS OF HUMAN CAPITAL MANAGEMENT TO ADDRESS

As demonstrated above, there have been a myriad of impacts on human capital in 2020 thanks to COVID-19. Companies within the P&C insurance industry must engage in thinking critically about these issues in order to retain a strong operational infrastructure and preserve their financial stability. With this in mind, three areas of human capital management for P&C insurance leadership should be top priority: communication, engaging civically, and investing wisely.

3.1 Communication

There is a business maxim that says: "when in doubt, over-communicate." Communication is vital to all business organizations. How much more so in times like these where uncertainty abounds! Lack of communication from leadership leaves employees to speculate about the future and creates insecurity and the stress described above. Taking a pro-active approach to communicate with employees builds trust and provides people with insight about potential changes on the horizon. Preparing mentally for a change is much easier than having it invade with no warning. Even when management does not know exactly what to say, P&C industry executives will benefit from communication with those in their charge. In doing so, the leader is displaying compassion and demonstrating that the manager is thinking about the consequences for those who are affected by any actions undertaken. Beyond individual conversations with employees, some practical suggestions might include employee listening initiatives, gathering opinions from various individuals throughout an organization, and company town-hall-style events (via Zoom of course!).

3.2 Civic Engagement

The circumstances presented by this pandemic require that business leaders engage civically with local and state jurisdictions to stay abreast of changes to the regulatory landscape. Executives and actuaries in the P&C insurance industry are attempting to project the future. However, no one could have forecasted what 2020 has thrown at us. The widespread nature of COVID- 19 has been unprecedented in recent times. Many of the organizations in the P&C insurance industry are spread across several geographic areas. This has complicated operational responses and makes it difficult to know what a proper outlook on business in the future should be. Regulators may implement laws affecting how the P&C insurance industry adjudicates claims or interacts with its own employees. Civic engagement will enable employers to stay informed so that they can provide the training and education their employees need.

3.3 Investment in Employees

There is no doubt that the financial impact on business due to COVID-19 has been immense. Since March, there have been almost daily headlines detailing workers being laid off as companies across the nation have had to lay off thousands upon thousands of people. At the end of September, insurance giant Allstate indicated it would lay off 3,800 employees. Financial strain can have a significant impact on the morale and psyche of the team that are so important to the success of an insurer. Challenging circumstances like these put pressure on management to invest wisely to preserve human capital. Offering personal financial management education could be a relatively inexpensive way to assist. Consideration of health-based incentives could protect long-term healthcare costs. It may be necessary to find creative teambuilding activities to help alleviate employees' fear that they are next to be laid off. Looking for options to increase productivity in a way that will increase return on investment should be considered as well. There is no doubt that balancing the obligations of an insurer during this time is difficult. Insurance executives who figure out a way to maintain financial strength while also prudently investing in their employees will benefit over the long-term.

4. REFERENCES

[1] W. Kenton, "Human Capital", *Investopedia*, https://www.investopedia.com/terms/h/humancapital.asp, **September 4, 2020**

[2] American Psychological Association, "Stress in the Time of COVID-19", Volume 1, https://www.apa.org/news/press/releases/stress/2020/stress-in-america-covid.pdf, **May 2020**

[3] K. Mayer, "HRE's number of the day: Coronavirus stress", https://hrexecutive.com/hres-number-of-the-day-coronavirus-stress/, **April 14, 2020**

[4] Geoffrey H. Tison, MD, MPH, Robert Avram, MD, MSc, Peter Kuhar, BS, Sean Abreau, MSc, Greg M. Marcus, MD, MAS, Mark J. Pletcher, MD, MPH, Jeffrey E. Olgin, MD, "Worldwide Effect of COVID-19 on Physical Activity: A Descriptive Study", https://www.acpjournals.org/doi/10.7326/M20-2665, **November 3, 2020**

[5] National Institutes of Health, "Higher daily step count linked with lower all-cause mortality", https://www.nih.gov/news-events/news-releases/higher-daily-step-count-linked-lower-all-cause-mortality, **March 24, 2020**

[6] Pulse, "Working-from-Home Perceptions", https://fluentpulse.com/covid-19-working-from-home/

Biography of the Author

Mark Maenche is a Consulting Actuary at Risk International Actuarial Consulting in Charleston, SC where he leads projects involving both pricing and reserving. In his role at Risk, he also employs his passion utilizing technology to drive efficiency in the production of analyses for clients. He participates on the CAS Candidate Liaison Committee. He holds a degree in Actuarial Science from the University of Illinois at Urbana Champaign and is an Associate of the Casualty Actuarial Society and a Member of the American Academy of Actuaries. He also holds the Certified Insurance Counselor and Certified Risk Manager designation from The National Alliance for Insurance Education and Research.

COVID-19 Models for Third- and Fourth-Party General Liability and Directors & Officers

Julie Menken, ACAS

1. INTRODUCTION

The highly contagious coronavirus, SARS-COV-2, and resulting respiratory illness COVID-19 has resulted in more than 152.5 million confirmed cases and nearly 3.2 million deaths across 237 countries as of May 3, 2021. While the ultimate severity of the pandemic remains uncertain, the current litigation environment in the United States indicates the possibility of mass litigation resulting from introduction of the virus via travel, failure to contain the virus and prevent spread of disease. Pandemic severity here is defined as the total damage from the virus which includes the number of U.S. deaths, the number of severe cases resulting in hospitalization, and ultimate healthcare costs of those severe cases. Such an event challenges (re)insurers to quantify and manage their exposure to 3rd and fourth-party pandemic losses. Exposure-based pandemic scenarios assist in this endeavor by stipulating mass litigation events that could arise from the pandemic and estimate the potential economic damages by company and industry. The following essay is meant to serve as commentary on how pandemic liability models are built and the factors considered regarding line of business specific COVID-19 liability models.

1.1 Scenarios Context

The basis of SARS-COV-2 scenarios requires building an epidemiological and exposure model which estimates the number of fatalities and hospitalizations under varying assumptions of the ultimate severity of the pandemic. A key estimate, the total number of COVID-19 fatalities, will depend on the success of public health measures in controlling the spread of the virus, the emergence of treatments for COVID-19, and how quickly a vaccine can be developed and deployed. As of the writing of this essay, the FDA has approved emergency use of three vaccines, however, the effectiveness of the vaccine to lower pandemic severity depends on public trust and willingness to get vaccinated, a significant public health hurdle to overcome. Pandemic liability models should consider a variety of data sources to build the basis for its scenarios including CDC, IHME, and WHO data sources, among other sources of information to project number of deaths and COVID-19-specific healthcare costs by age and industry. Projected U.S. fatalities under each pandemic severity level of the liability model should be allocated among residents of long-term care facilities, workers in essential industries, workers in non-essential industries, and non-workers. Industry and company specific damages vary by estimated employee count and risk level.

1.2 Objective

Models for COVID-19 liability will differ based on third-party versus fourth-party coverages due to differences in how these mass litigation events have typically played out in the U.S. court system. In this essay, we discuss both and the precedent set by asbestos mass litigation and how this may factor into future pandemic liability lawsuits.

2. THIRD-PARTY TAKE-HOME LIABILITY SCENARIOS

A key segment of third-party pandemic liability models considers mass litigation scenarios regarding liability for "take-home coronavirus" exposures whereby workers contract the virus in workplace and then infect their family members and/or co-habitants. Household members sickened with COVID-19 then sue the employers of infected workers for damages.

While the first official case of COVID-19 in the U.S. was detected in Washington state on January 19, 2020, epidemiologists believe the virus was circulating in major cities weeks before. As the number of cases and deaths accelerated in late February and early March, city and state officials enacted "stay-at-home" orders which closed many businesses, banned public and private gatherings, and directed residents to remain home for all but essential travel. However, there was significant variation in how and when businesses responded to the pandemic. Some businesses put social distancing practices in place or closed entirely prior to the enactment of stay-at-home orders. Others continued to operate as normal until directed to close by public health authorities. Essential businesses ranging from grocery stores to construction sites to hospitals continued to operate through the stay-at-home orders.

Businesses that continue operations during the pandemic have potential to contribute to the spread of SARS-COV-2. Workers may contract the virus from their co-workers or customers. Workers who contract the virus in the course of employment may be able to claim Workers' Compensation benefits depending on the jurisdiction. Some workers could also file personal injury lawsuits against their employers if they believe their employers were negligent in exposing them to the virus.

The focus of 3rd party pandemic scenarios are on the emergence of lawsuits filed by individuals who alleged they contracted COVID-19 from a co-habitant who was infected with SARS-COV-2 at work. These take-home coronavirus plaintiffs argue that the employer was negligent in exposing the worker, and subsequently themselves, to the virus and sue for damages associated with COVID-19, including medical costs, lost wages, and pain and suffering. Under this scenario, take-home coronavirus lawsuits emerge in a wide range of industries including nursing homes, meatpacking, construction, public transportation, food service, and essential retail operations such as grocers and building supply stores. Individual lawsuits are consolidated in federal and state multidistrict litigations. Industry-by-industry, juries in bellwether cases hold defendants liable for take-home coronavirus plaintiffs' injuries and award damages. While a global settlement is not possible due to the diversity of defendants, a general case management strategy emerges that is employed widely to achieve efficient settlements for thousands of pending cases.

From this base scenario description, models may target three levels of pandemic severity and three levels of propensity to recover damages, considering the uncertainty vaccine rollouts and mask adherence to public health recommendations as well as the state-by-state legal considerations of duty. All scenarios include an estimate for the percent of COVID-19 hospitalizations and deaths which are attributable to take-home exposures. It is not known how likely it is that these injured individuals will file lawsuits and recover damages. Plaintiffs face difficulties in establishing that their infections resulted from a given employer's negligence and furthermore that the employer acted negligently and should be held liable. Workers' Compensation, health insurance, and possibly a government-administered essential worker victims' compensation fund offers alternative mechanisms for compensation that may dampen litigation claim rates. The uncertainty surrounding alternative compensation mechanisms is also reflected in the levels of propensity to recover.

Estimates identify over 200 industries with exposure to take-home coronavirus lawsuits. Scenario losses distribute losses among these industries in proportion to a qualitative measure of ease of virus spread and quantitative estimates of industry employment. Damages are distributed to companies within those industries in proportion to estimated employment. Curated and established portfolio modeling software would apply an insurance model to the pandemic liability scenarios and allow clients to run their general liability and umbrella portfolios against the scenarios to receive an estimate of their COVID take-home liability exposure.

In the current U.S. litigation environment, there is a non-insignificant probability that the United States court system will experience mass litigation from coronavirus injuries. Precedent for take-home exposure litigation was established with asbestos litigation whereby injured family members successfully sued the employer of the family member who brought home asbestos fibers on their clothes. Cases of asbestos-related disease from such second-hand exposure are well documented and some have resulted in successful litigation against the initial employer. Such take-home exposure can be used as precedent in 3rd-party take-home COVID-19 mass litigation lawsuits. The first take-home COVID-19 lawsuit was filed in May 2020 in Illinois. As of May 3, 2021, fifteen take-home coronavirus infection complaints have been filed, most alleging damages based on a negligence theory but some on a public nuisance theory, which has also been seen in asbestos cases. Given observations of these early cases building on each other as test cases, the expectation is that we will see more of these take-home liability lawsuits in the years to come.

3. FOURTH-PARTY LIABILITY SCENARIOS

The second segment of pandemic scenarios envisions state governments filing suit against airline carriers and cruise lines alleging their negligent actions resulted in introduction and spread of SARS-COV-2 in the United States. The resulting lawsuits in this scenario class seek to recover costs of treating COVID-19 patients under state Medicaid programs. The fourth-party scenarios will have the same underlying epidemiological and exposure model as the third-party scenarios. The difference here is in the defendants and the mechanisms of mass litigation.

The United States enacted travel restrictions to and from China on January 31 and to and from Europe on March 12, but hundreds of thousands of travelers entered the U.S. before international air travel was effectively suspended worldwide, allowing the virus to be widely introduced to major population centers in the US. State and local governments, compelled to mobilize a response to the pandemic in the absence of federal government action, have at the same witnessed steep declines in their tax revenues and their budgets have been hit hard by the coronavirus. The Tax Policy Center estimates state government revenue will decrease by $200 billion across fiscal years 2020 and 2021 relative to pre-pandemic forecasts. The decline in state government revenue has resulted in massive budget shortfalls that, in the absence of aid from the federal government, can only be addressed by reducing expenditures and raising additional revenue. At the same time, state governments are being forced to spend additional resources to expand emergency and hospital services to care for the hundreds of thousands of individuals who have been sickened by the virus.

With fourth-party pandemic liability, the model may contemplate a scenario by which state attorneys general file suit in state and federal courts seeking to recover the cost of treating COVID-19 patients by state Medicaid programs. The lawsuits name 60 air carriers and cruise lines who may have brought infected travelers into the United States between the time when the coronavirus was first known to be a major public health threat and when the United States effectively suspended international travel. These companies are alleged to have downplayed the seriousness of the health threat, continuing to operate "business as usual," and not taking precautions commensurate with the risk such as discouraging non-essential travel, aggressively screening passengers for signs of illness prior to boarding, mandating all passengers wear masks, limiting occupancy, and disinfecting all surfaces between flights.

Individual state lawsuits are consolidated in federal multidistrict litigation. Juries in bellwether trials are persuaded that defendants' disregard of the public health threat posed by coronavirus accelerated the introduction and spread of the virus in the United States resulting in greater public health expenditures than would have materialized had these companies acted with an appropriate level of care in the early days of the pandemic. Juries hold defendants liable for the costs they imposed on public health insurance programs awarding several state plaintiffs hundreds of millions of dollars in compensatory and punitive damages. A global structured settlement modelled after the Tobacco Master Settlement Agreement is eventually reached between all plaintiffs and defendants with settlements allocated across defendants according to estimates of the number of passengers disembarking with coronavirus.

Fourth-party litigation operates by different standards of liability than do third-party suits. As such, the lawsuits do not consider individual cases against essential businesses, but rather, lawsuits against major players who can be deemed responsible for introducing the virus to the U.S. by failing to heed warnings in early days. In this way, the defendants in these scenarios will be lower in number with a greater percentage of loss allocation to each. Allocation to the estimated defendants in fourth-party scenarios will be based on extent of travel and date of first actions take to stop the spread of the virus via travel restrictions. Fourth party pandemic liability models may use international passenger statistics for cruise ships and airlines as a basis for potential loss allocation to insureds. Fourth-party scenarios also should contemplate the capacity to bear loss and market cap of the potential defendants. In particular, 4th party liability scenarios may also consider the issue of international conventions which limit liability for airline carriers.[3]

4. D&O AND OTHER CONSIDERATIONS

A third segment of pandemic liability scenarios should also consider is event-driven D&O, which have been increasing in recent years. Event-driven securities litigation involves operational risks which are not appropriately disclosed to investors and then a later event reveals the issue and causes the stock value to tumble. Mass tort litigation can serve as the vehicle for revealing the undisclosed operational risk, and therefore the same scenarios described for general liability can simultaneously trigger D&O events. As a result, there is significant risk of cross-line clash that casualty insurers face, and it is necessary to prepare companies by understanding their exposure not only to COVID-19 general liability but also the potential for them to see litigation across lines of business.

When mass torts cause a clash of underlying general liability and D&O losses, D&O is the "caboose" of the liability train. Using recent examples outside of COVID-19 of mass torts against opioids manufacturers and distributors, 3M and other PFAS manufacturers and users, Monsanto and J&J, event-based securities litigation can be modeled for potential COVID GL-D&O clash risk. Models for such risk will quantify the risk by estimating the size of the resulting event, the potential range of shareholder damages, and the range of shareholder recoveries. To estimate shareholder damages models would estimate share price movements and reflect known historical patterns of shareholder recoveries. While the details of a COVID19 D&O model may differ from the underlying GL model, the underlying pandemic severity assumptions are consistent and clash between lines of business contemplated. Considering the litigation uncertainty and pandemic severity uncertainty, COVID19 liability models help provide estimates for the extent of insurer exposure both within and across lines of business.

5. REFERENCES

[1] World Health Organization COVID-19 site: https://www.who.int/emergencies/diseases/novel-coronavirus-2019

[2] IHME projects over 5 million COVID-19 deaths globally and nearly 599 thousand U.S. COVID-19 deaths August 1, 2021, as of this writing. These numbers consider a rapid vaccine rollout. https://covid19.healthdata.org/global?view=total-deaths&tab=trend

[3] Potential COVID-19 Liability Facing Air Carriers https://www.jdsupra.com/legalnews/potential-covid-19-liability-facing-air-34753/

Abbreviations and notations

CDC, Centers for Disease Control
D&O, directors and officers
GL, general liability
IHME, Institute for Health Metrics and Evaluation
PFAS, per- and polyfluoroalkyl substances
WHO, World Health Organization

Biography of the Author

Julie Menken, ACAS is a Senior Client Engagement Manager and Actuary at Praedicat, a liability emerging risk analytics and casualty catastrophe modeling company for casualty insurers and global industrial companies. Praedicat's emerging risk framework, liability scenarios and stochastic loss model makes emerging risk actionable across its lifecycle, helping companies to better identify liabilities early, track the risks and take action as they mature, and defend claims if litigation emerges. Praedicat is creating the technology for a growing and sustainable casualty market.

Praedicat was established in 2021, is based in Los Angeles, California and has offices in New York and London. Praedicat successfully participated in Lloyd's Lab Cohort 3 and 5, partnered with SOMPO to win an SMA Underwriting Innovation award in 2019, and has made the InsurTech Impact 25 of 2020 list by Oxbow Partners as one of the top-25 most promising insurance technology solution providers. Praedicat's ultimate aim is to deliver the science around health and environmental risks to businesses, driving smarter decisions that make the world cleaner, safe and healthier.

The Secret Life of Trend—Including Other Models for Estimating Trend and Credibility for Trend

Joseph Boor, FCAS, Ph.D., CERA

Abstract

Actuaries use loglinear trend regularly. However, there are several aspects of trend that are not common knowledge among actuaries. Three key issues are: the loglinear model for trend is not the only model for trend; it is affected by uncertainty arising from the loss development process; and, there is not much in the actuarial literature creating formal mathematics-based credibility formulas for trend. Two alternatives involving the effects of random drift on expected losses in addition to the effects of trend are presented. One including the point-by-point error associated with regression, and one without it, are presented. Further, there are alternate algorithms for computing trend. Trend estimation is discussed in all three contexts and using those alternate algorithms. An evaluation of the impact of loss development uncertainty on trend is provided. Corresponding credibility formulas for trend are provided as well.

1 Introduction

The loglinear trend process is well established. But for certain situations, the standard algorithm may be unwieldy. For example, when using the linear regression approach, determining the uncertainty in the trend formula engendered by the underlying uncertainty the loss development may be challenging. Further, the regression formula for determining trend using the logarithms of the data points, involves a very specific model of trend. In effect, it assumes that there is a constant trend effecting each year to year step, but the data points are affected by error values with a common variance σ^2. That may not reflect the reality of the data. So, within this paper, optimum trend models are presented, alternate algorithms for calculating the trend under each formula are presented, and some situations in which the alternate algorithms are useful to actuaries are presented.

2 Models of Trend and the Corresponding Formulas for Estimating Trend

Many actuaries use regression-based loglinear trend. However, as mentioned earlier, it has several features that may not be obvious, and it contains an exposure to loss development uncertainty.

2.1 A Comment on the "Log" Part of "Loglinear" Trend

Most of the common models of loss cost inflation/trend recognize that inflationary forces are better modeled by nonlinear models. In effect, because trend is believe to compound, inflationary models with assumptions like those used in compound interest are typically used. So, the cost level of the losses at some time may be modeled by some $C(t) = \exp(a + bt)$ (where, elsewhere in this paper, a may not be constant from time to time). Therefore, it is helpful to work with the natural logarithm of the cost level $\ln(C(t)) = a + bt$ to get data that has a simplified, linear, character.

Of course, in most cases, the actual values of $C(t)$ for various t's are not available. Rather the (often) annual historical loss severities "$f(t)$", loss frequencies "$n(t)$", pure premiums "$\pi(t)$", or annual loss ratios "$LR(t)$", present in the data are what is available. They can be expected to differ from the true underlying expected severities, frequencies, etc. by some error amount.

The next issue to resolve would be deciding which probability distribution best reflects that range of that error. The Central Limit Theorem provides a rationale for the normal distribution, since a number of claim values are added together, then divided by the number of exposures. However, the normal distribution produces negative as well as positive values, so it is not always a realistic model for a trend driven by any sort of loss cost inflation. The lognormal distribution, on the other hand, is based on a geometric rather than arithmetic version of the Central Limit Theorem. Essentially, the lognormal error scenario is presumed to involve a large number of (positive) error terms that are multiplied by, not added to, the true costs. Therefore, the lognormal approach does not produce negative values. Further, the lognormal error is more consistent with trending and loss development calculations. So it is used (via it's σ parameter) throughout the remainder of this paper. Additionally, the geometric/multiplicative lognormal error approach underlies the loglinear trend that is used so often by so many actuaries[1]. Recognizing all those considerations, the lognormal modelof the trend will be used in much of this analysis.

Since the lognormal generates random variables that are exponential functions of normal distributions, one is left with a generalized trending equation, say for pure premium, of $\pi(t) = \exp(a+bt)$, where the constant $\exp(a)$ and the growth factor $\exp(b)$ may each be subject to error and other randomness. Taking logarithms of the $\pi(t)$, etc. values produces a much-more tractable linear-type model $\ln(\pi(t)) = a+bt$. So, throughout the remainder of this paper the focus will be on the simpler, linear type, algorithms. Thus, although the ultimate goal is to estimate the trend, the majority of this paper will cover the slope of the regression line.

2.2 The Loglinear Trend Model: Constant Trend with Process Error

Of course, the loglinear trend approach is the workhorse trend estimation method used by casualty actuaries. Nevertheless, there are aspects of it that may be of interest.

2.2.1 The Basic Approach Used in Loglinear Trend

To clarify the discussion in subsection 2.1, the classic loglinear trend involves the following underlying analysis:

1. There is an underlying constant geometric trend factor $1 + T$ that causes the underlying expected pure premiums ($E[\pi(t)]$'s), expected severities, or whatever else is being analyzed to grow exponentially ($E[\pi(t + 1)] = (1 + T)E[\pi(t)]$) as t increases.

2. Of course, if the expected losses above from item 1 were known, determining the trend would be trivial. However, in practice, few trend datasets are that perfect. So, it is then assumed

[1] Admittedly, though, that is more of an endorsement by the crowd than the result of statistical principles.

that even though the expected losses have a perfect pattern, the data are subject to some error[2] that acts independently, but with a common variance, on each value. So, each of the π's is considered to be the true data $E[\pi(t)]$ multiplied by one of a series of equal independent lognormal distributions "$M(t)$", all with mean "1" (unity) and having identical coefficient of variation "v". So, each historic loss ratio, etc. value is $\pi(t) = M(t)E[\pi(t)]$.

3. The logarithmic transformation reduces the results of number 2 to a set of $\ln(M(t)E[\pi(t)]) = \ln(\pi(0)) + t\ln(1+T) + \ln(M(t))$. That is a constant, a slope multiplied by time t, and a set of identical normally distributed process error[3] terms, each with mean zero. After estimating the optimum values of the slope and constant with regression, the projection of any future point $\pi(t+s)$ may be found using that line formula and computing the exponential function on the results.

4. That format involves fitting a constant and a slope so that the constant plus the product of a slope and time minimizes the sum of squared differences between the fitted values at the various times and the actual data points[4]. It amounts to using regression to determine $y = a + tb$ given the historical data points[5] (the $y_1, y_2, ..., y_k$) and independent time variables $t_1, t_2, ..., t_k$.

5. To simplify the notation, the remainder of this section will simply focus on expressing the regression using t's and y's.

Then, the linear model with process error assumes that the logarithms of the expected cost level, the $E[y_i]$'s, indeed lie on a line and follow

$$E[y_i] = a + bt_i. \tag{1}$$

But, the available historical data is different from the true expected cost levels, and each regression data point y_i differs from the underlying $E[y_i]$ by some normally distributed error, $err(i) = E[y_i] - y_i$. The $err(i)$'s, per the model, all are expected to be independent of one another and have the same variance, σ^2. So, $y_i = E[y_i] - err(i)$ for all i, and

$$y_i = a + bt_i - err(i);\ \text{for all } i. \tag{2}$$

Of course, much of the basics of linear regression are part of the basic education of casualty actuaries. But the material above is presented in order to provide complete clarity on the exact assumptions underlying loglinear regression. This will also set up the approach to be used for other models of trend.

---

[2]For example, this might be process or parameter error.

[3]In the mathematics,this is referred to as "“observation error”. This is potentially a broad definition of process error that would include any independent error variables sharing a common variance and a common mean of zero.

[4]Minimizing the sum of squared differences would, per the mathematics of the normal distribution, amount to the maximum likelihood estimator of the constant $\ln(\pi(0))$ and the slope $\ln(1+T)$

[5]Normally, n would be used for the number of points, but it was already used to denote the number of points, but it has already been used to represent the frequency. So, "k" will denote the the number of historical data points used in the regression underlying the trend.

2.2.2 The "Weights" Assigned to the Regression Data Points

The next step is to look at the "objective" function that the trend estimate is designed to minimize. The underlying likelihood that a set of points $y_1, ..., y_k$ are generated from a given a and b, is a constant, times the exponential function, of the negative of

$$\sum_{i=1}^{k} \frac{(y_i - a - bt_i)^2}{\sigma^2}. \tag{3}$$

So, minimizing that sum of squared differences maximizes the probability that the historical y's could arise from expected costs that follow the trend line.

Unsurprisingly, this devolves to finding an a and b that minimize a sum of squared errors. But, surprisingly, it does not depend on σ^2. One may begin with the covariance formula for the slope, and not yet specify that the times used in the regression are regular annual, quarterly, etc. Then, focusing on the averages, $\bar{y}$ of the y_i's and $\bar{t}$ of the t_i's, the slope may be written as

$$b = \frac{\sum_{i=1}^{k}(t_i - \bar{t})(y_i - \bar{y})}{\sum_{i=1}^{k}(t_i - \bar{t})^2} = \frac{\sum_{i=1}^{k}(t_i - \bar{t})y_i}{\sum_{i=1}^{k}(t_i - \bar{t})^2} = \sum_{i=1}^{k} \frac{t_i - \bar{t}}{\sum_{j=1}^{k}(t_j - \bar{t})^2} y_i \tag{4}$$

(noting that the constant $\bar{y}$ in the first term is multiplied by some values that add to zero.)

Therefore, the slope b is just a linear combination of the values in the regression data. Further, theory of sums of series indicates that the denominator is equal to $\frac{k^3-k}{12}$. The values from below the mean $\bar{t}$ of time are negative, the others are positive. The simpler expression for the slope is

$$b = \sum_{i=1}^{k} \frac{t_i - \bar{t}}{\frac{k^3-k}{12}} y_i. \tag{5}$$

Of course, in practice the t_i's are consecutive years, consecutive quarters, or something similar. Hence, it makes sense to focus on examples using consecutive and evenly spaced times.

As an example, lets say the values from 2011, 2012, 2013, 2014, and 2015 are to be used to estimate the slope. $\bar{t}$ is clearly 2013=year 3. The value of the denominator is $\frac{k^3-k}{12}$ is $(5^3 - 5)/12 = 10$, treating 2011 as year one. Further, the "k" "weights" starting from that of 2011, are -2/10, -1/10, 0, 1/10, 2/10 = -.2, -.1, 0, .1, .2. Note the linear progression stemming from equation (4), and the symmetry up to a minus sign. Both are general characteristics when the times are year to year, quarter, to quarter, etc. without breaks.

It is easy to see that the midpoint $\bar{t}$ of the numbers $t_i = i = 1, 2, ..., k$ is $\frac{k+1}{2}$. So, the "weights" for computing the slope from annual data may simply be stated as

$$12\frac{i - \frac{k+1}{2}}{k^3 - k} = 6\frac{2i - k - 1}{k^3 - k}. \tag{6}$$

Thus, the slope is really just a difference of "weighted sums"[6] of the y values. One may also note that the values $12\frac{i-\frac{k+1}{2}}{k^3-k}$, $12\frac{j-\frac{k+1}{2}}{k^3-k}$ have constant denominators, so the points further from

[6]The slope is not a weighted average of the values. It rather a difference between weighted sums. One may see that the "weights" sum to zero, since the value $i - (k+1)/2$ at i is always the negative of the value $(k+1)/2 - i$ on the opposite side at $(k+1) - ii - (k+1)/2$ at the time on the opposite side $(k+1) - i - (k+1)/2$ of the center at $(k+1)/2$.

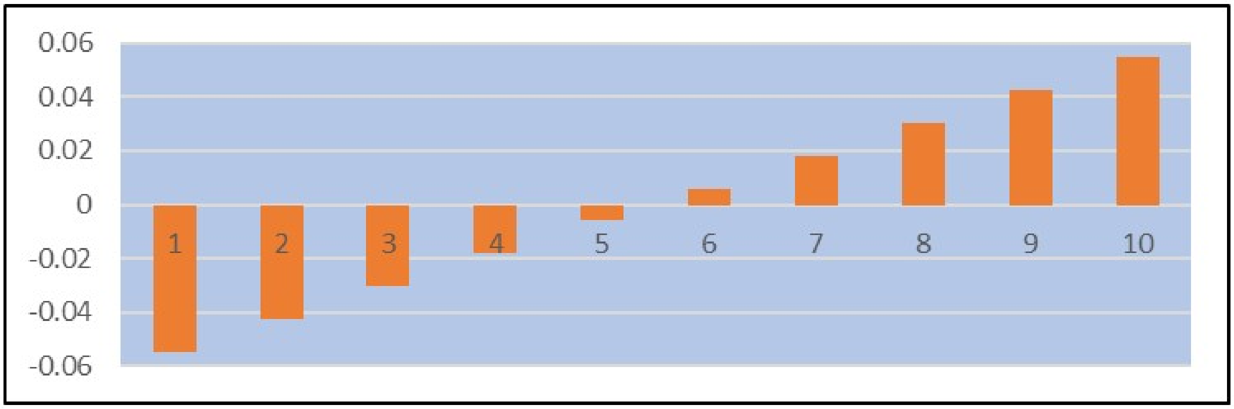


Figure 1: "Weights" (Multipliers) Underlying All Ten Point Regressions

the center, where $|i - \frac{k+1}{2}|$ is larger, have greater influence. For further illustration, consider the "weights" underlying ten point regression, They are shown in Figure 2.2.2.

One may readily see that the endpoints receive the largest weight. Also, the "weights" are symmetric, excepting that, as before, the earlier "before trend has happened" values are negatives, while the later "after trend had happened" values are positive. Thus, the trend is based on the difference between those groups of points. However, having the point-by-point values allows for a more advanced analysis, as will be shown in the next section.

2.2.3 The Impact of Loss Development Uncertainty

The last section sets the stage for an analysis of the impact of loss development uncertainty on trend. Consider a regression slope computed using five points. The middle point has zero weight, the first two have negative weight, and points four and five have positive weight. Arguably, either all the negative or all the positive values can be thought of as determining the slope. Between points four and five, the last, and the most developed[7] data point is point five, which has two-thirds of the positive weight. For a ten point slope, the last point has 36% of the positive weight.

If one has perspective on the development uncertainty in the data to be trended, and the uncertainties in the various points are statistically independent, one may estimate the variance of the slope due to development uncertainty. If an estimated value has a standard deviation of the possible ultimate losses of ϕ, then the the standard deviation of its logarithm may be estimated by $\gamma = \phi \times \frac{dlog}{dx}(\mu) = \frac{\phi}{\mu}$, or the coefficient of variation of the distribution of possible ultimate losses. Given that the various points' uncertainties are independent, one need only multiply the resulting values (squared, for variance) by the squares of the weights in equation (9) to get the consequential variance of the trend estimate across possible actual values of the ultimate losses.

In many cases, the variances of the logarithms of the ultimate losses, etc. may be directly estimated (and be statistically independent), perhaps by using the approach in Hayne 1985. Then the total variance of the error in the slope estimate due to loss development uncertainty is simply the sum of the variances due to development at the various years, etc., each multiplied by the square of the corresponding "weight"

[7] Exactly how much development is involved of course depends on the line of business, perhaps the class of business, etc.

$$36\frac{\sum_{i=1}^{k}\gamma_i^2(2i-k-1)^2}{(k^3-k)^2} \tag{7}$$

(where each γ_i^2 is the variance of the logarithm of the i^{th} data point.)

As one may see, this could sometimes be quite substantial . On the other hand, though, the result above does not reflect the full error variance associated with the slope estimate from regression. The regression result is also only an estimate, thus the error it makes predicting the slope is also part of the total error variance of the resulting slope.

If one has a fairly good handle on the variance of the process error, the results may be improved some by switching from the standard regression to "weighted regression". Weighted regression in this case may be illustrated by the goal or objective function it seeks to minimize. Standard regression minimizes the squared differences between the points on the line and the data values. Weighted regression, weights are assigned to each of the squared differences. They correspond to the total variance (process and loss development) affecting each point. Thus one would seek[8]

$$\sum_{i=1}^{k}\frac{(a+bi-y_i)^2}{\sigma^2+\gamma_i^2} = min. \tag{8}$$

Another alternative is to use calendar year trend, which requires no loss development. However, one must weigh that against its susceptibility to say, a claims department's decision to close a large percentage of their inventory in one of the calendar years (and how much that might distort the trend) and the fact that the data is from somewhat older accident/report years.

2.2.4 The Weights (Yes Weights) of the Year-to-Year Differences in the Regression Data

Section 2.2.2 provided the "weights" for determining the slope as a linear combination of the data points. The next step is to show that the estimated slope from consecutive times $t = 1, 2, ..., k$ is a weighted average of the year-to-year increases $y_{i+1} - y_i$. The first step in computing the weights involves noting that, when k is odd, the weight w_1 for $y_2 - y_1$ must match (be the negative of) the subsection 2.2.2 point-by-point type weight for y_1 of

$$-6\frac{k-1}{k^3-k}. \tag{9}$$

A moment's review of the sums will show that the weight w_i applied to $y_i - y_{i-1}$ must equal the "weight" from section 2.2.2 for the point y_i in isolation, or $\left(\text{i.e. } 6\frac{2i-k+1}{k^3-k}\right)$ less w_{i-1}. The provides a point-by-point formula for the w_i's.

That formula for the w_i's may be solved, and the resulting weights for the one year slopes are

$$w_i = 6\frac{i(k-i)}{k^3-k} \text{ for each } y_{i+1} - y_i. \tag{10}$$

The corresponding weights for the nine one year slopes in ten point regression are shown in Figure 2.2.4.

[8] As information, extensions to some popular spreadsheet software packages that perform this calculation are available at present if one does not wish to use a goal seek solution routine to compute this.

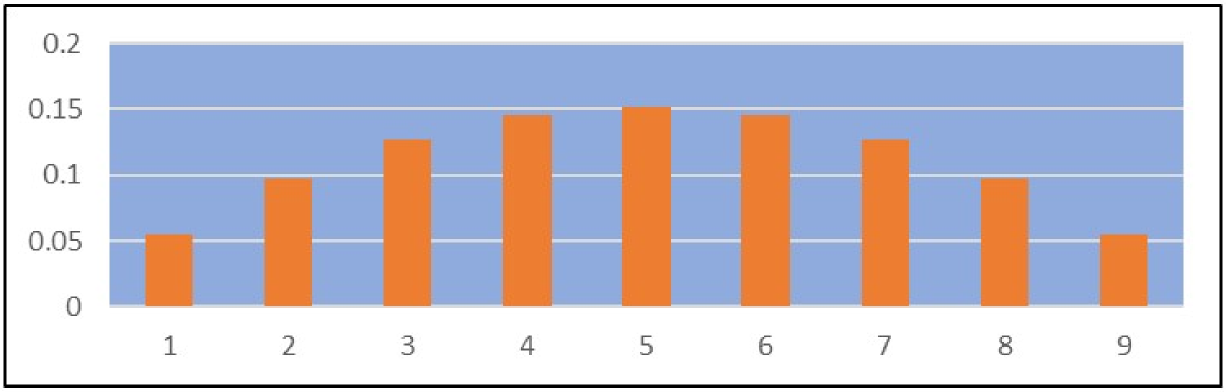


Figure 2: Common (All Ten Point Regressions) Weights for the Nine One Year Slopes in the Data

These weights (w_i's) have a very important property—they sum to unity. Thus, the projected linear trend (slope) b is really just a weighted average of the year-to-year slopes in the data. So the original external projected trend ratio $T+1 = \exp(b)$ will be a geometric average of the year-to-year growth values in that data. The same weights will be used, but they will represent exponents for the various year-to-year growth values within the geometric average. Further, one should note that although the weights for individual points y_i are larger as one moves away from the center of the experience period, the weights for the $y_{i+1} - y_i$'s are larger near the center of the period.

2.2.5 Summary of the Results for Regression

In conclusion, the regression slope (logarithm of the trend value) may be expressed as a difference between weighted sums of the loss, etc. values, or as a weighted average of the year-to-year changes. That leads to an estimate of the effect of loss development uncertainty on the fitted slope. Further, it is well known that the optimum prediction under regression to some period $k + j$ is a straight average (identical weights of $1/k$) of the y values, plus the calculated slope times the number of years from the mean time $\bar{t}$ to the future period. As shown above, for this annual data, both the beginning point and the slope to future periods are weighted values of the y's. So any linear projection to some future period $k + j$ may be expressed as a weighted sum/linear function of the y's.

$$est(y_{k+j}) = \sum_{i=1}^{k} \left[\frac{1}{k} + \left(\frac{k+1}{2} + j \right) \frac{6(2i - k + 1)}{k^3 - k} \right] \hat{y}_i. \tag{11}$$

(Essentially, the $\frac{y_i}{k}$ sum to the mean of the y_i's. The $\frac{k+1}{2}$ trends from the mean time associated with the mean of the y_i's to the time associated with the last data point. Lastly, the j term moves it to the future time period desired for the projection.)

Overall, one may see that the standard loglinear trend algorithm is based on computing a straight average for the starting point and a weighted average for the trend.

2.3 The Trend with Random Drift Model: Varying Trend but no Process Error

The loglinear approach deals with situations where the true underlying expected loss values for each year are not completely known, but the underlying trend is constant. The trend with random drift case involves perfectly known expected loss values. However, in addition to the trend, those expected losses "drift" in a random way.

This is really another use for a model that is widely employed by some other financial service providers. Use of that model, geometric Brownian motion, to reflect changing costs is not new. It is commonly used in the investment community as a model of risk-adjusted stock price evolution. It has already been used in the actuarial world in Boor 1993 and McNichols and Rizzo 2012.

Basically, it assumes that the cost, etc. levels $C(t)$ are affected by a constant trend (T as always). However, the cost level is also buffeted by constant but random changes, so that from year to year its logarithm is changed by a random selection from a normal distribution (in addition to the long-term slope). The effects of these changes are cumulative in that all the prior changes are embedded in each value. In the transformed distribution $\ln(C(t))$ values at time s and time t differ not only by the logarithm of the trend, but also by some value from a normal distribution with some variance parameter δ^2. It may be written

$$\ln(C(t)) - \ln(C(s)) = (t-s)\ln(1+T) + (t-s)\delta N(0,1) \tag{12}$$

(where $N(0,1)$, in a slight abuse of notation (for clarity), represents a sample from the standard normal distribution).

One may describe $\delta N(0,1)$ as "random drift". Since it is linear now, it is associated with the slope of a line rather than with the compounding trend in the original trend data. So, this is "slope with random drift" rather than "trend with random drift". Notably in this linear case, in each set of intervals (s,t) and (u,v) that do not overlap (other than at the endpoints), the samples from the normal distribution are independent.

Figure 2.3 contains an example of what the expected loss ratio (in this case,without process variance, also the historical losses). might look like in this scenario.

That means that the yearly slopes $y_2-y_1, y_3-y_2, ..., y_k-y_{k-1}, (y_i = \ln(C(i))$ are all independent samples of the slope $b = \ln(1+T)$. Since there are $k-1$ samples, the slope estimate using the year-to-year changes is clearly

$$est(b) = \frac{1}{k-1}\sum_{i=2}^{k} y_i - y_{i-1} \tag{13}$$

One may note that any i not on the top or bottom of the range, is included in both $y_{i+1} - y_i$ and $y_i - y_{i-1}$. So most of the terms cancel, leaving a slope estimate for these historical points under slope with random drift of

$$est(b) = \frac{1}{k-1}(y_k - y_1). \tag{14}$$

In effect, the weights for the points are $-\frac{1}{k-1}$ for y_1, $\frac{1}{k-1}$ for y_k, and zero for the other points.

To finish the linear portion of this analysis, one may note that slope with random drift is well-known to be "Markov" or "memoryless". That means that for times after the last data point at k, y_k alone is the best point for making future predictions. $y_1, y_2, ...y_{k-1}$ would only be useful when y_k and perhaps additional values are not known. So, whereas predicting future values using regression in equation (11) involved using middle of the time values at the starting point, slope with random

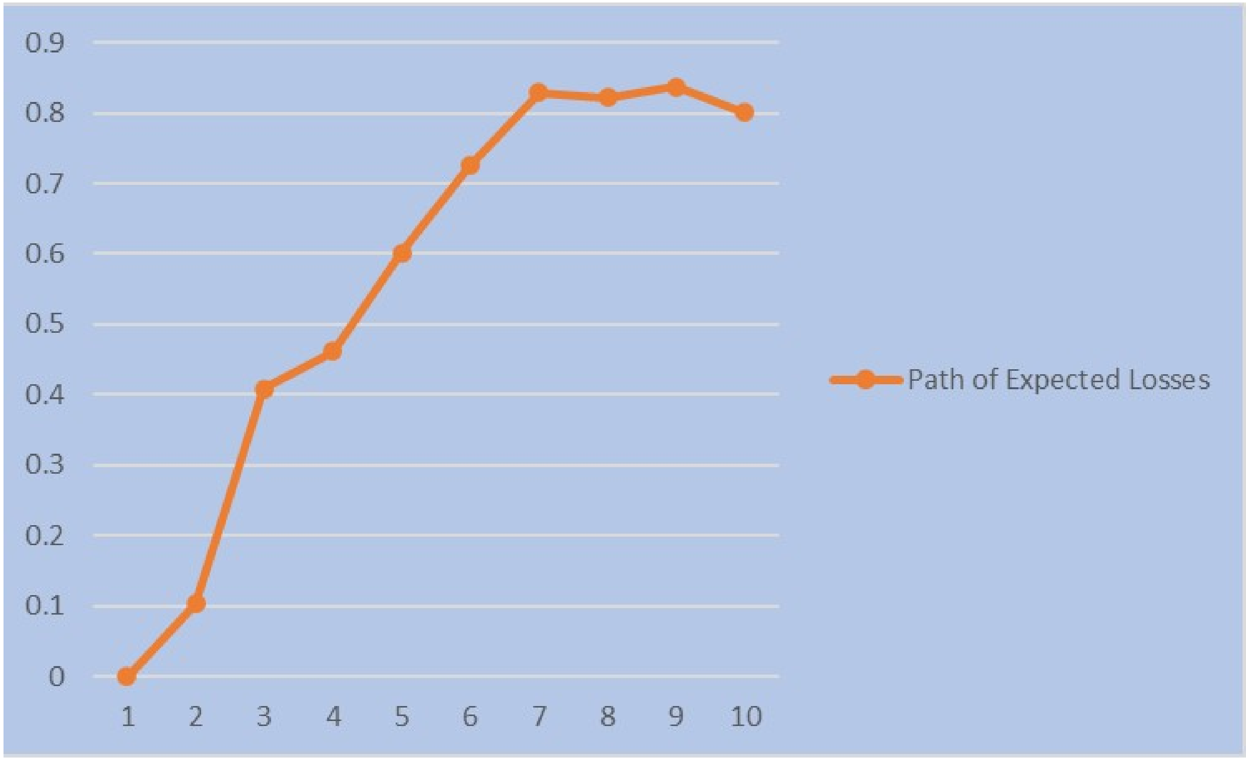


Figure 3: Example of Expected Loss Ratios Buffeted by Random Drift

drift predictions begin with the most recent data point y_k. So the linear slope with random drift estimate for y at time $k+j$ is

$$est(y_{k+j}) = y_k + \frac{j}{k-1}(y_k - y_1) = \frac{j+k-1}{k-1}y_k - \frac{j}{k-1}y_1 \tag{15}$$

Unlike the regression system, translating the slope with random drift formulas back to the needed trend with random drift values ($C(i)$'s) is easy. Equation (14), based on the difference between the last point and the first point becomes a ratio, And the multiplier $\frac{1}{k-1}$ becomes an exponent, generating the $k-1^{st}$ root of $1+T = \left(\frac{C(k)}{C(1)}\right)^{\frac{1}{k-1}}$. And per the Markov property $est\left(C(t+j)\right) = C(k) \times \left(\frac{C(k)}{C(1)}\right)^{\frac{j}{k-1}}$.

2.4 Trend with Both Random Drift and Process Error

Contrary to the assumptions of the last two sections, sometimes trend is influenced both by random drift and process error. This section presents a model to use in such a situation.

2.4.1 Explanation of the Model

The previous models each include a core assumption that could sometimes be an unrealistic . Often trend data from very large datasets that would seem to be susceptible to the trend with random drift view of subsection 2.3 appear to be a little different from what the theory would suggest. For example, consider the consumer price index data in Table 1, where one would expect little process error. Process error would create a situation where very large increases or decreases in the trend could come from more extreme errors. In that case, a large decrease in the annual trend would be followed by a large increase and vice versa. One cannot determine conclusively from the data, but the changes in the trend from 2008 to 2009, 2009 to 2010, and 2010 to 2011 do suggest that some process error (perhaps arising from the data collection process) is present.

Table 1: Year-to Year Trend Rates in Consumer Price Index (All Urban Consumers

Value at 12/31 of Year	CPI	Change in CPI
2006	210.800	
2007	210.036	-0.36 %
2008	210.228	0.09 %
2009	215.949	2.72 %
2010	219.179	1.50 %
2011	225.612	2.94 %
2012	229.601	1.77 %
2013	233.049	1.50 %
2014	234.812	0.76 %
2015	236.565	0.75 %

As one may see, the trend rates in the CPI data are fairly volatile. Further, it is possible that the dynamics of the consumer price index involve even more complexity. Therefore it is reasonable to question whether or not the loglinear trend model really captures the structure of the data it is applied to. The concern with trend with random drift is more direct. One may also see that the trend with random drift assumption that none of the points contain process error of any sort, is suboptimal for datasets subject to process risk.

Therefore, an approach that recognizes both the process, etc. risk that makes the data points imperfect representations of the underlying costs and also accommodates drift-type volatility from year to year is needed. The approach begins with the trend with random drift process consistent with

$$E\left[\ln\left(C(i+1)\right)\right] = E\left[\ln\left(C(i)\right)\right] + \ln(1+T) + \delta N(0,1) \tag{16}$$

(with the notation abusing $N(0,1)$ terms representing independent standard normal samples for each of the various intervals $(i, i+1)$).

Thus, $E\left[\ln\left(C(i+1)\right)\right]$ follows the slope with random drift paradigm. But, rather than the pure random drift involved in subsection 2.3, this scenario also includes the process error included in subsection 2.2. Specifically, one may state that

$$y_i = E\left[\ln\left(C(i)\right)\right] + \sigma N(0,1) \tag{17}$$

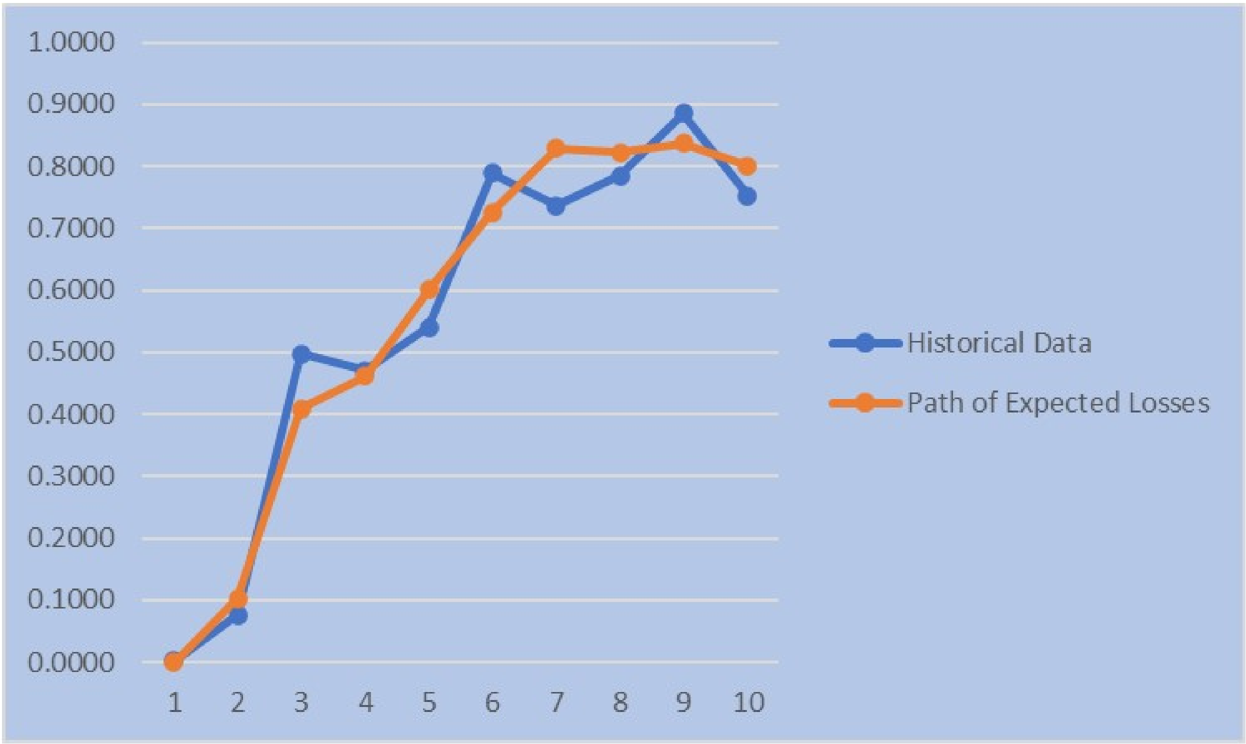


Figure 4: Example of Random Drift/Process Variance Combination

(with the $N(0,1)$ terms independent among the various indices $i = 1, 2, ..., k$).

Figure 2.4.1 illustrates this scenario. Both the expected loss ratios that are subject to random drift and the historical loss ratios, that combine random drift and process variance, are shown.

With this understanding, one my start to develop methods for estimating the slope from the historical data. That will be pursued in the next section.

2.4.2 Finding the Key Variances

Considering the presence of both types of volatility, the goal is to find the a and b that are most consistent with the data. To do that, one must first define an error function or objective function to minimize. The most obvious approach would be to take a page from the playbook of the other two situations and seek the slope that is consistent with the lowest possible variance. However, in this case there are actually two variances, both the process variance from subsection 2.2 and the drift variance from subsection 2.3. So one must consider what combination of those should be minimized.

For illustration, note that σ and δ (actually σ^2 and δ^2) may be estimated using historical data. Per Boor 2015[9], once the slope is removed σ^2 and δ^2 may be computed using

$$\frac{E\left[\sum_{i=1}^{k-1}(y_{i+1} - y_i)^2 - (y_k - y_1)^2\right]}{2(k-2)} = \sigma^2, \tag{18}$$

[9] To facilitate its use, be aware that in the referenced paper the "y_i's' were labeled as "S_i"s to limit conflicts among variable names.

and

$$\frac{E\left[(k-1)(y_k-y_1)^2-\sum_{i=1}^{k-1}(y_{i+1}-y_i)^2-(y_k-y_1)^2\right]}{(k-1)(k-2)}=\delta^2. \tag{19}$$

where each y_j is the value of the linear (generally, log-transformed) value for the j^{th} year, month, etc.

However, when one attempts to simultaneously estimate the slope, process error variance σ^2, and drift variance δ^2, the problem tends to become too unwieldy to perform reliably, at least per per a few methods employed by the author. Therefore, one may suggest using the trend with random drift approach when the values appear to be fairly compact around a curve, the approach of this section when there is a similar non-linear appearance, but the values are not compact around the curve, and the regression approach otherwise, at least as a starting point.

2.4.3 Estimating the Slope

On the other hand, if reasonable estimates of σ^2 and δ^2, may be made, then it as least possible to provide an estimate for the slope. The idea involves estimating the underlying "expected loss"[10]/slope with random drift/no process error path, then use the standard slope with random drift estimate of the slope of the expected values, [(most recent point) $-$ (first point)]/$(k-1)$, of the slope from subsection 2.3.

To do so, it is helpful to define the best approximation point-by-point as $e_1, e_2, ..$ etc. These are to be computed/estimated, along with the variance of each of them around the unknown true expected losses/cost level (the τ_i's) at each time. For example, at the first data value y_1, the only information available is the value y_1, so that would be the estimate e_1 of the first point. Its variance around the true value[11] on the underlying path of the expected values of losses would be σ^2, which can be set as the initial value τ_1^2.

For the second value along the path, we have two estimators, e_1+b and y_2. The expected prediction variance of y_2, with respect to the value of the underlying expected losses would logically be it's variance from the those losses, or σ^2. The expected squared prediction error generated by e_1+b would be the squared error inherent in e_1, or the process variance σ^2, plus the inherent volatility as one moves from year to year along the path, δ^2. Since the two may generally be thought to be independent, the variance of the error between the estimate e_1+b and the true value along the path is $\delta^2+\sigma^2$.

That begins the iteration. Since the expected squared error e_1 makes predicting the expected value is $\tau_1^2=\sigma^2$, and the drift along the path is independent of the process error, the error e_1+b makes (where b is the currently unknown slope) in predicting the second true point on the path adds one year of random drift to make $\tau_1^2+\delta^2$ the error variance of e_1+b in predicting the second point on the path. The error variance of y_2 would be σ^2. A formula from best estimate credibility (see Boor 1992) indicates that for these two independent estimators, the best estimate results from

[10]This could also be expected frequency, severity, etc

[11]The language is key here. In this case, the mean of e_1 is equal to the actual underlying first point on the path the expected losses follow as they drift. However, if they did not match, the expected squared error predicting the initial point on the path would have to include the squared difference between the mean of e_1 and the mean of that initial point along with the variance of e_1. Considering that all the distributions used in this section are presumed to be unbiased and independent, per Boor 1993, it should not be an issue. However, it is mentioned for completeness and clarity.

weighting each one by the expected squared prediction error of the other[12]. Therefore, the (best) estimated value of the second point on the path the expected losses underlying the data actually followed is

$$e_2 = \frac{\sigma^2(e_1 + b) + (\tau_1^2 + \delta^2)y_2}{\tau_1^2 + \delta^2 + \sigma^2}. \tag{20}$$

Since the two components are (clearly) independent, the variance of the result above is just the result of multiplying the variances of the two by the scalar multipliers (like credibilities). A little algebra results in in a formula for the error variance of e_2 of

$$\tau_2 = \frac{\sigma^2 \times (\tau_1^2 + \delta^2)}{\tau_1^2 + \delta^2 + \sigma^2}. \tag{21}$$

Those may be generalized into recursive formulas for the e's and τ's

$$e_{i+1} = \frac{\sigma^2(e_i + b) + (\tau_1^2 + \delta^2)y_{i+1}}{\tau_i^2 + \delta^2 + \sigma^2}. \tag{22}$$

$$\tau_{i+1} = \frac{\sigma^2 \times (\tau_i^2 + \delta^2)}{\tau_i^2 + \delta^2 + \sigma^2}. \tag{23}$$

Of course, since that formula assumes that one already knows the slope, it is not directly useful for estimating the slope. However, it may be used indirectly. In the regression model, the slope is set so that squared residuals between the fitted line and the actual points are minimized. So, it would be logical to seek the value of b for which the iterations of equations (22) and (23) generate the least squared residuals between the best estimate e_i's and the actual points (y_i's). Table 2 illustrates the process when a 10% exponential trend is accompanied by process error corresponding to $\sigma^2 = .005$ after the logarithmic transform and drift variance $\delta^2 = .002$, also in the logarithmic transformed data. Of course, the exponential/trend with random drift data is first converted to a linear system using logarithms so that all the variables are on a linear basis. Then the calculations mentioned earlier are carried out in Table 2. Lastly, the spreadsheet software searches for the value in medium gray of the slope that minimizes the sum of squared differences between the data points and the best estimate of the points along the underlying path. That, after conversion to loglinear trend, forms the trend estimate.

The historical values and the estimated points on the path are shown in Figure 2.4.3.

One could question whether the fairly high accuracy (estimate of 9.85% vs. an actual 10.0%) of this method is due to the larger slope predominating over the two variances. Therefore, the same calculations were done for a trend rate of 3% in Table 3.

The sums of squared differences in Table 3 match Table 2 because the data to be analyzed was the same, up to the slope, in both examples. However, notice that even the estimate of the lower 3% trend was very, very good. For full perspective, though, in the experience of the author if the data looks like it complies with the regression assumptions, typically this only provides a marginal improvement in accuracy over the regression estimate. Further, as the number of data points increases, the various methods are more and more prone to yield similar trend estimates. This method may be more useful when the process variance is very small, when the pattern of the data points contains a curve, hump, or something else inconsistent with the regression assumptions,

[12] If the reader is so inclined, one may verify that the formula works in this instance by using Bayesian methods.

Table 2: Estimation of Underlying Slope With Both Random Drift and Process Error When Actual Value (10%) is Larger than the Standard Deviations

Constants:

A. $\sigma^2 =$	0.0050,	$\sigma =$	.0706	(considered known)
B. $\delta^2 =$	0.0020,	$\delta =$	.0447	(considered known)
Exponential Trend =	10.00%,	Slope of Logs =	9.53%	(both to be found)

Year	(1) Simulated Loss Ratio	(2) "S" Natural Log of Loss Ratio	(3)=[Prior (6)] "e" Current Point Est of Expected Loss Level	(4)=[Prior (7) Incoming Variance "τ"	(5)=(4)+B. Drift Variance to Next e	(6)={[(3)+b]A.+(1)(5)} /{A. +(5)} Next e	(7) = (5)×A. /(A.+(5) Outgoing Value of τ	(8)=((2)-(3))^2 Difference Data Point S and e
1	101.3 %	0.0128	0.0128	0.0050	0.0070	0.1034	0.0029	0.00000
2	110.4 %	0.0987	0.1034	0.0029	0.0049	0.1925	0.0025	0.00002
3	120.6 %	0.1876	0.1925	0.0025	0.0045	0.3128	0.0024	0.00002
4	140.0 %	0.3365	0.3128	0.0024	0.0044	0.4383	0.0023	0.00056
5	159.3 %	0.4657	0.4383	0.0023	0.0043	0.4822	0.0023	0.00075
6	155.1 %	0.4389	0.4822	0.0023	0.0043	0.6342	0.0023	0.00188
7	198.2 %	0.6843	0.6342	0.0023	0.0043	0.6619	0.0023	0.00251
8	183.1 %	0.6047	0.6619	0.0023	0.0043	0.7690	0.0023	0.00328
9	218.2 %	0.7803	0.7690	0.0023	0.0043	0.8587	0.0023	0.00013
10	235.2 %	0.8551	0.8587	0.0023	0.0043			0.00001

b=Estimated Slope of Logs= 9.40%
T=Est Loglinear Trend= 9.85%
Sum of Differences Between Loss Level Path and Data Points= .00916

or there are fewer data points. These alternate views are also better suited to situations where one expects that the underlying trend changed significantly during the time period of the data.

In conclusion, while there is a workable formula for estimating the variance structure given knowledge of the trend, there is also a workable formula for estimating the trend given the variance structure. However, the author is not aware of any good approach to estimate both simultaneously. Nevertheless, in certain situations, investing the time needed to execute this method can yield better accuracy in the slope estimate.

3 Limited Fluctuation Credibility for Trend

Now that formulas that relate the trend calculations to individual points are available (at least for regression and pure trend with random drift), it is possible to develop credibility formulas that are designed specifically for trend. The two versions for limited fluctuation credibility follow.

3.1 What Would a (Limited Fluctuation) Credibility Formula for Trend Look Like?

When considering credibility for trend, it is relevant to begin with the core goal of the given credibility process. Although actuaries typically think of limited fluctuation credibility in terms of claim counts, it is really about the rate, trend, etc. not changing too much unless the data clearly indicate that a given change is needed. For example, for the common 1082 standard, the objective is to not allow pure randomness in the data to arbitrarily change rates by more than 5% (up or

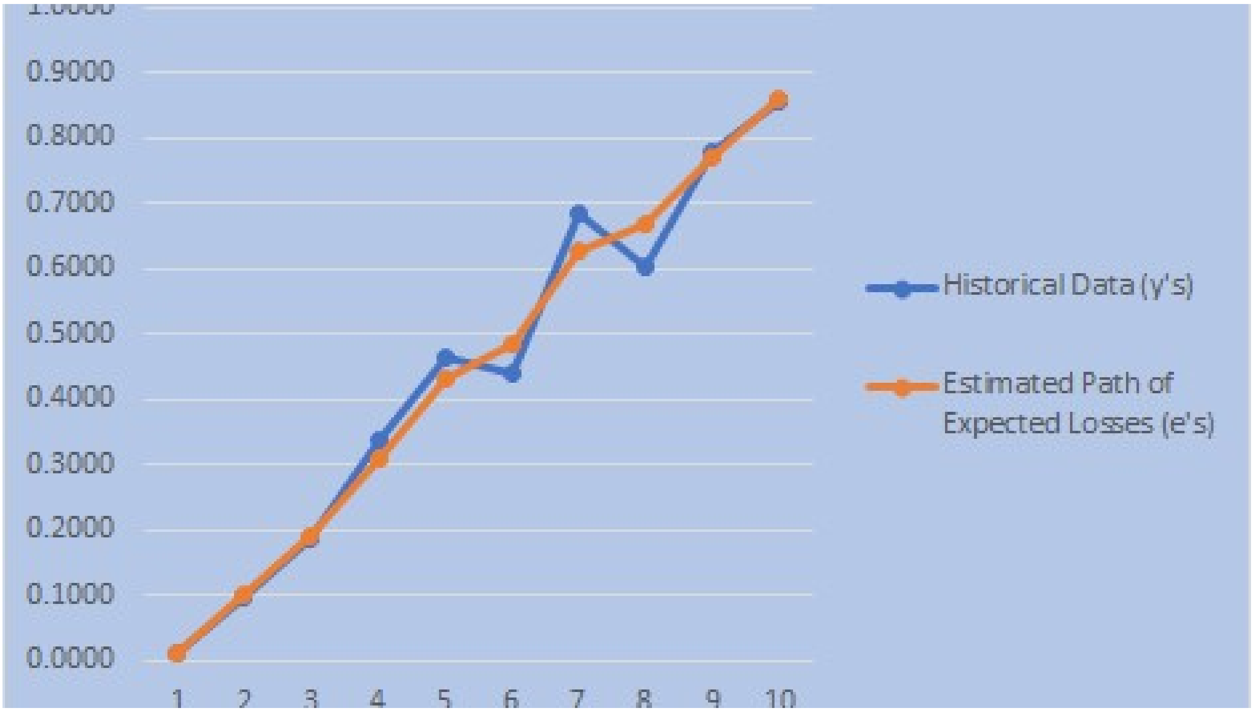


Figure 5: Historical Values and Estimated Path of Expected Loss Ratio From Table 2

down), unless the data indicate such a change is needed. Further, since any amount of loss can conceivably happen, "the data indicate such a change is needed" is defined as "There is a 10% (or less) chance that the credibility-adjusted result will randomly create more than a 5% change." Those requirements are not based on claim counts, claim counts are merely a convenient way to compute the credibility.

So, in the case of limited fluctuation credibility, the stated goal is that probability that random chance causes losses to exceed some threshold ± R is limited to some suitably low probability p. So, the goal is to find (or estimate) the pth and 1−pth percentiles of the distribution of possible trends. Then one may appropriately throttle the distribution with a credibility factor Z so that $Z \times F^{-1}_{trend}(\text{p}) \geq -\text{R}$ and $Z \times F^{-1}_{trend}(1-\text{p}) \leq \text{R}$, where $F^{-1}_{trend}(p)$ is the pth percentile of the distribution of possible changes in trend.

3.2 Limited Fluctuation Credibility with Regression-Based Trend

Subsection 3.1 allows one the opportunity to define a general credibility formula for the regression slope. Such a formula would depend on the variances, not on claim counts. Specifically, it is well known that approximately 90% of the probability in the normal distribution is within 1.645 standard deviations (up or down) of the mean (note that for purposes of considering pure randomness, the mean would be "no change"). So if Z is set to be $(maximum\ acceptable\ change\ in\ slope) \div (1.645 \times CV(linear\ slope))$(where CV denotes the coefficient of variation of the slope estimate, the standard deviation divided by the mean), the criteria underlying the 1082 standard (90% chance of no change beyond the maximum acceptable change) will be fulfilled by credibility weighting the slope with some very reliable ancillary data. However, that does require one to know the variance

Table 3: Estimation of Underlying Slope With Both Random Drift and Process Error When Actual Value (3%) is Smaller than the Standard Deviations

Constants:					
A. $\sigma^2 =$	0.0050,	$\sigma =$	.0706	(considered known)
B. $\delta^2 =$	0.0020,	$\delta =$	.0447	(considered known)
Exponential Trend =	3.00%,	Slope of Logs =	2.96%	(both to be found)

Year	(1) Simulated Loss Ratio	(2) "S" Natural Log of Loss Ratio	(3)=[Prior (6)] "e" Current Point Est of Expected Level	(4)=[Prior (7) Incoming Variance "τ"	(5)=(4)+B. Drift Variance to Next e	(6)={[(3)+b]A.+(1)(5)} /{A. +(5)} Next e	(7) = (5)×A. /(A.+(5) Outgoing Value of τ	(8)=((2)-(3))^2 Difference Data Point S and e
1	1.013 %	0.0128	0.0128	0.0050	0.0070	0.0363	0.0029	0.00000
2	1.034 %	0.0330	0.0363	0.0029	0.0049	0.0603	0.0025	0.00001
3	1.058 %	0.0561	0.0603	0.0025	0.0045	0.1124	0.0024	0.00002
4	1.149 %	0.1393	0.1124	0.0024	0.0044	0.1695	0.0023	0.00072
5	1.225 %	0.2027	0.1695	0.0023	0.0043	0.1570	0.0023	0.00110
6	1.116 %	0.1101	0.1570	0.0023	0.0043	0.2336	0.0023	0.00220
7	1.336 %	0.2898	0.2336	0.0023	0.0043	0.2074	0.0023	0.00315
8	1.155 %	0.1445	0.2074	0.0023	0.0043	0.2442	0.0023	0.00397
9	1.290 %	0.2543	0.2442	0.0023	0.0043	0.2682	0.0023	0.00010
10	1.301 %	0.2633	0.2682	0.0023	0.0043			0.00002

b=Estimated Slope of Logs= 2.82% Sum of Differences Between Between Loss Level Path and Data Points= .00916
T=Est Loglinear Trend= 2.86%

of the fitted slope around the true slope.

It should then be clear that the main challenge in determining credibility for the regression slope that underlines the trend is finding the percentiles of the distribution of possible slopes. However, since the regression assumptions suggest a normal distribution, only the standard deviation is actually needed. Thankfully, a standard statistic is available to identify the standard deviation. Using the special additional regression option available in a common spreadsheet software package, one may compute the key statistic needed. After first taking logarithms of the CPI data in Table 1, the spreadsheet option produces (as approximately excerpted from the regression output).

Table 4: Excerpt from Supplementary Information Spreadsheet Software Provided in Loglinear Regression of Table 1 Data

ANOVA				
	df	*SS,*	*MS*	*F*
Regression	1	0.01965	0.01965	173.0653602
Residual	8	0.00091	0.00011	
Total	9	0.02056		
	Coefficients	Standard Error,	t Statistic	p value
Intercept	-25.62411	2.35860	-10.86411	4.55693E-06
X Variable	0.01543	0.00117	13.15543	1.06096E-06

The values in gray are the the mean (fitted) slope and the standard error (error standard deviation) of the regression slope. Thus, the 1082-equivalent credibility calculation for the linear regression would be $5\%/(1.645 * (.00117/.01543)) = 40\%$

However, the credibility of the final trend T of the actual CPI values is slightly different. To estimate the trend in the original CPI data, one must invert the logarithm and convert the trend factor $1 + T = \exp(slope)$ to the trend rate $T = \exp(slope) - 1$. That affects the relative error (which is supposed to, under the 1082 standard, be 5% or less) through the magnification or shrinkage generated by the exponential function, as well as the denominator used in computing the relative error. The magnification or shrinkage multiplier would be generated by the derivative of $\exp(x) - 1$, or $\exp(x)$ at $x = .01543$. The value is 1.01555. Then, since $T = \exp(.01543) - 1 = .01555$, the 5% relative error allowed in determining the slope translates to relative error of $5\% \times 1.0155 \times .01543/.01555 \approx 5.04\%$. So the 5% threshold still essentially holds in this example. However, it is relevant to complete this final check to be certain that the credibility fulfills its function appropriately.

3.3 Limited Fluctuation Credibility with Trend with Random Drift

Just as in loglinear trend, the key to limited fluctuation credibility for trend with random drift trend lies in computing the variance, specifically, the variance between the computed slope and the actual slope. In this case, the process is much more straightforward. In say, ten years, of data there are nine year-to-year slopes. It is not difficult to calculate the variance of those slopes. Then, since the estimate of the trend with random drift is simply the average of those nine slopes, all one need do is divide the variance of the individual slopes by nine. That estimates the variance of the error in the slope estimate, and the remaining process mirrors that used in subsection 3.2.

3.4 Usefulness of Limited Fluctuation Credibility for Trend

Of course, once it is computed, limited fluctuation credibility can be used in a wide variety of situations. This can be used when the complement of credibility benchmark is countrywide trend for the line of business, or when it is the trend in last year's rate review. The flexibility and (comparative) ease of computing this are offset somewhat by the fact that it does not result in the most accurate estimate of trend. The methods of the next section will focus on accuracy, but consequently the formulas are less robust.

4 Best Estimate Credibility for Loglinear Regression

As noted in Boor 1992, best estimate credibility depends not just on how well the data predicts the loss costs, it also depends on how well (or poorly) the complement of credibility benchmark predicts the loss costs. It further depends on whether or not the prediction errors generated by the data and the benchmark are correlated. If one is using countrywide trend as a benchmark, one might expect the errors to be uncorrelated. However, if one is using the trend generated last year as a benchmark, one might expect substantial correlation. Therefore, the two approaches are analyzed in separate sections.

4.1 Best Estimate Credibility with External Benchmark Data

Per Boor 1992, best estimate credibility is a function of the error each statistic (dataset) makes in predicting the underlying quantity being estimated (in this case, the slope underlying the trend).

For two predictors that make uncorrelated errors, the credibility weight of one statistic is proportional to the squared error generated by the other statistic. One might expect that, say, trend computed from countrywide data would make prediction errors that are largely uncorrelated with those generated by the trend data of a small state. In general, most benchmarks tend to make prediction errors that are unrelated to the those made using a dataset with lesser volume. So that simpler, no covariance, formula would apply.

The next step that is required is to estimate the expected squared errors associated with the two predictors (the slope computed from the subject data and the benchmark slope). However, as discussed in subsection 3.2, that may be begun, for both the target data and the benchmark, by using the standard errors of the two slopes obtained in the regression process.

For the subject data, that suffices to produce an estimate of the squared error. However, the squared error the benchmark makes in estimating the subject slope solely requires another term, because the benchmark and the subject data simply have different underlying slopes. Therefore, the squared error the benchmark makes is more than just variance from the regression fit. That amount, at least in this characterization, is a constant bias[13] rather than variance. It does contribute to the squared error, though. It is not hard to see that the expected squared error is equal to the variance of the predictor plus the square of that bias.

Now, one may only know the actual bias by knowing the underlying slopes that are being estimated. However, one could use the difference between the slope of the subject data and the slope of the benchmark data to estimate the bias. Then the credibility of the slope in the subject data is:

$$\begin{aligned} &Z(\text{subject data}) = \\ &\qquad \left[\text{standard error}^2(\text{benchmark}) + (\text{difference of slopes})^2\right] \\ &/ \left[\text{standard error}^2(\text{subject data}) + \text{standard error}^2(\text{benchmark}) + (\text{difference of slopes})^2\right]. \end{aligned}$$

For example, When the Table 1 data is loglinearly regressed, the slope is 0.01543 and the standard error is 0.012. If one could then identify a benchmark to supplement this (CPI) data, and it had a slope of .017 and standard error of .003, the the credibility of the CPI data would be $\frac{.003^2+(.01543-.017)^2}{.012^2+.003^2+(.01543-.017)^2} = 74\%$. Thus, given the regression output, this is not a challenging calculation.

It should be apparent that this general approach will also work with the slope with random drift using the variance of year-to-year changes formula from subsection 3.3. Details are not provided as the formula should also be apparent.

Of course, at this point, with either trend scenario, a credibility weighted estimate of the slope is produced, but what is actually needed is the exponential-based trend. The procedure for converting the expected squared error in the slope to that in the exponential trend has already been discussed in subsection 3.2. Alternately, one may simply perform the credibility calculation on the regression slopes, and then convert the result to a trend by applying the exponential function to the slope and subtracting unity.

[13]Technically "bias", although it does not have the strong negative connotation associated with bias that distorts the results—as this improves the results.

4.2 Best Estimate Credibility When Updating Loglinear Regression

In updating trend, the complement of credibility is either last year's trend or last year's slope. It seems logical to use the same credibility for each. In contrast to the uncorrelated nature of the errors in external benchmarks with internal data, last year's slope uses much of the same actual data as the subject data. So, then Boor 1992 indicates (after some algebra) that the optimum credibility for the slope is

$$\begin{aligned}
&Z(\text{subject data}) = \\
&\left[\text{standard error}^2(\text{old slope}) + (\text{difference of slopes})^2 - \text{covariance(new slope,old)}\right] \\
&\Big/ \Big[\text{standard error}^2(\text{new slope}) + \text{standard error}^2(\text{old slope}) + (\text{difference of slopes})^2 \\
&-2 \times \text{covariance(new slope,old)}\Big] .
\end{aligned} \tag{24}$$

The next issue to be resolved, then, is that of estimating the covariance between this year's slope and last year's slope. The first tool lies in some additional output of the regression calculated earlier using spreadsheet software. The data in Table 5 also comes from the regression on the Table 1 CPI data.

Table 5: Additional Excerpt from Supplementary Information Spreadsheet Software Provides in Loglinear Regression of Table 1 Data

Multiple R	0.97766
R Square	0.95582
Adjusted R Square	0.95029
Standard Error	0.01066
Observations	10

The standard error in gray is the standard deviation of the "residuals", or differences between the fitted curve and the data points. Combining this with the weights from equation (9) in subsection 2.2, that are effectively used in computing the slope, one may see that the variance of each (independent) point/residual and weight combination is equal to the square of the weight times the standard error squared. To show that this works note this alternate computation of the previously provided (Table 4) standard error of the slope.

$$\sqrt{.01066^2 \times \sum_{i=1}^{k} \left(6\frac{2i-k-1}{k^3-k}\right)^2} = .00117. \tag{25}$$

The goal, though, is to compute the covariance of this year's slope with last year's slope. That may be done in a similar fashion. It is not hard to see that the first year in the current trend period was the second year in last year's trend period, and so on. Then, if we further define r_{new} to be the standard error in the new regression and r_{old} to be that of the prior regression, one may obtain (after some algebra) the following formula for the covariance

$$
\begin{aligned}
Cov(\text{new slope,old slope}) &= \sum_{i=1}^{k-1} r_{new} r_{old} \times 6\frac{2i-k-1}{k^3-k} \times 6\frac{2i-k+1}{k^3-k}. \\
&= r_{new} r_{old} \times 12\frac{k-3}{k(k^3-k)}.
\end{aligned}
\tag{26}
$$

Note that this is essentially a constant, identical across all updates and data used in k period regression, times the product of the two standard errors.

One may also note that this does not completely resolve the updating credibility problem. It works when the complement of credibility term is last year's slope, but not when it is the credibility weighted average of several prior years that updating would have generated last year. However, that problem is fairly complex. Per H. Gerber and D. Jones 1975, when these sort of covariances exist, successive updates potentially, perhaps likely, require changes in the credibility mix of older years to be truly optimal. So, each credibility-weighted average could easily contain a very different set of weights, with no guiding formula to simplify constructing the covariance. Likely, there is still some view of what is optimal that would accommodate this particular situation. Hopefully, this provides a first step. Further, in context the formulas of this subsection could conceivably be used to provide a proxy for the full updating credibility.

5 Summary

A detailed analysis of the trend, and associated linear slope, calculations was presented. Three alternate scenarios for the underlying process driving the trend were presented, along with some guidance for estimating the trend in each case. How often they produce materially different trends is not known at present, but might represent an opportunity for further analysis. Lastly, using details of the analyses of the trend process in this article, credibility formulas on both a limited fluctuation and best estimate basis were provided. Those formulas focused primarily or conventional regression trend, but form a template for the trend with random drift as well.

References

J. Boor. Credibility based on accuracy. *Proceedings of the Casualty Actuarial Society*, 79:165–185, 1992.

J. Boor. A stochastic approach to trend and credibility. *Casualty Actuarial Society Forum,Special Ratemaking Edition*, pages 341–400, 1993.

J. Boor. The credibility of the overall rate indication: making the theory work. *Variance*, 9:167–186, 2015.

H. Gerber and D. Jones. Credibility formulas of the updating type. *Transactions of the Society of Actuaries*, 27:31–46, 1975.

R. Hayne. An estimate of statistical variation in development factor methods. *Proceedings of the Casualty Actuarial Society*, 72:25–43, 1985.

J. McNichols and J. Rizzo. Stochastic gbm methods for modeling market prices. *Casualty Actuarial Society Forum,Summer Edition*, pages 21–18, 2012.

The Pareto-Gamma Mixture

Greg McNulty, FCAS

Abstract: In this paper we will review some established properties and derive some new properties of a Pareto distribution with fixed scale whose unknown shape parameter is Gamma distributed. Namely:

- that Gamma is a conjugate prior to the Pareto distribution
- the formula for the posterior parameters of the Gamma given observed data
- a closed form for the CDF of the Pareto-Gamma mixture
- that the mean and all higher moments of the distribution are infinite
- a formula for the moments of the limited expected value of a random variable following this distribution
- the intractability of the closed form

Keywords: Pareto distribution, Gamma distribution, Conjugate prior, Bayesian statistics, Reinsurance pricing

1. INTRODUCTION

Our company recently performed a study comparing the exposure based large loss models for a long-tailed line of business across a number of reinsurance cedents. The industry curves indicated losses excess of a common threshold followed a Pareto distribution with the scale parameter varying by cedent. Fitting a distribution to the sample of shape parameter values showed that they approximately followed a Gamma probability distribution.

Hoping to benefit from the simple formulas of a conjugate prior relationship for a Bayesian model (and avoid more a difficult programming exercise using R) we found in the *Wikipedia* entry for "conjugate prior" [1] that Gamma was indeed conjugate prior to the Pareto. However, unlike for almost every other such pair of distributions, the closed form of the posterior predictive distribution was not given. We wondered if this was because it was not available or did not exist.

It turns out the closed form does exist, and this paper will show a derivation relying at times on other well-established facts rather than providing full mathematical proofs. We will also explain why the mean and all higher moments of the distribution are infinite. Finally, we will derive the formula for powers of the limited expected value of x, although we will find them to be of limited usefulness.

Please note, different texts and tools will use different parameterizations of the gamma distribution. This document attempts to be self-consistent using the definition in Section 2.2.

2. CONJUGATE PRIOR RELATIONSHIP

Proofs will be supplied for each step, but the reasoning is as follows: the Pareto distribution is "log-Exponential"; Gamma is conjugate prior to the Exponential distribution; the conjugate prior relationship is preserved under the log transformation; therefore Gamma is conjugate Prior to "log-Exponential", aka the Pareto distribution.

2.1 Pareto is "Log-Exponential"

To see this, let's just write the PDF of a random variable whose log is exponentially distributed:

$$f(\ln(x)) \sim e^{-\lambda * \ln(x)} = x^{-\lambda}$$

To be a little more precise, we could start with the CDF of the Pareto:

$$F(x) = 1 - \left(\frac{x}{\theta}\right)^{-\lambda}; x > \theta$$

Now we can transform x into a random variable supported on $[0, \infty)$ by letting $y = \ln(x/\theta)$. Then we would have:

$$F(y) = 1 - (e^y)^{-\alpha} = 1 - e^{-\alpha y}$$

This is the CDF of the exponential distribution, meaning a Pareto distributed random variable's scaled natural log is exponentially distributed. We also see here the relationship between the parameters of the distributions, namely being equal under the correct parameterizations.

2.2 Gamma is Conjugate Prior to Exponential

For this section we will use Bayes' Theorem:

$$p(\theta|x) \sim p(x|\theta) * p(\theta)$$

where x represents the data and θ represents the model parameters. We use the following parameterizations of the Exponential and Gamma distributions:

$$f(x; \lambda) = \lambda e^{-\lambda x}$$

$$\Gamma(x; \alpha, \beta) \sim x^{\alpha-1} e^{-\beta x}$$

In this case conjugate prior would mean that if *x* is Exponentially distributed, and $p(\lambda)$ is Gamma distributed, then $p(\lambda|x)$ is also Gamma distributed. Note that by *x* we mean a set of *n* observations of the random variable, which could also be denoted $\{x_i\}$. Then we have:

$$p(x|\theta) = \prod_i \lambda e^{-\lambda x_i}$$

Using Bayes' Theorem:

$$p(\lambda|x) \sim \left(\prod_i \lambda e^{-\lambda x_i} \right) * \lambda^{\alpha-1} e^{-\beta\lambda}$$

$$= \lambda^{\alpha+n-1} * e^{-(\beta+\sum x_i)\lambda}$$

$$\sim \Gamma(\lambda; \alpha + n, \beta + \sum x_i)$$

So, the posterior distribution of the Exponential parameter is again Gamma distributed, and we also have expressions for the posterior parameters of the Gamma distribution.

2.3 Conjugate Prior Relationship Preserved Under Logarithm

Now we can show that Gamma is a conjugate prior to the Pareto distribution. Suppose *x* is Pareto distributed:

$$F(x) = 1 - \left(\frac{x}{\theta}\right)^{-\lambda} ; x > \theta$$

Further suppose the Pareto α is Gamma distributed:

$$\Gamma(\lambda; \alpha, \beta) \sim \lambda^{\alpha-1} e^{-\beta\lambda}$$

Typically, the Pareto parameter shape parameter is called "alpha", but in order to avoid confusion with the Gamma parameter we choose a different letter. Let's again apply the transformation $y = \ln(x/\theta)$. Then we see *y* is Exponentially distributed according to:

$$G(y) = 1 - e^{-\lambda y}$$

Now suppose we observe a set $\{x_i\}$. This is equivalent to a set $\{y_i\} = \{\ln(x_i/\theta)\}$. Because Gamma is conjugate prior to the Exponential, we know the posterior distribution is:

$$p(\lambda|\{y_i\}) \sim \Gamma(\lambda; \alpha + n, \beta + \sum y_i)$$

But the λ of the Exponential is the same as the λ of the Pareto. We know that the Bayesian posterior distribution is Gamma and we can rewrite the above equation as:

$$p(\lambda|\{x_i\}) \sim \Gamma(\lambda; \alpha + n, \beta + \sum \ln(x_i/\theta))$$

Which means that Gamma is conjugate prior to the Pareto.

3. CLOSED FORM OF THE MIXTURE

We use the same logarithm transformation of a Pareto distributed variable into an Exponentially distributed variable to access the simpler derivation of the closed form for the Gamma-Exponential mixture.

3.1 Gamma-Exponential Mixture

Suppose we have that:

$$f(x; \lambda) = \lambda e^{-\lambda x}$$

$$\mathrm{g}(\lambda; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$$

In other words, *x* is Exponentially distributed, with the rate parameter itself a variable which is Gamma distributed. Can we obtain a closed form for this mixture, i.e. a formula for *f(x)* unconditional on λ that does not contain integrals?

The derivation below was posted on StackExchange [2] courtesy of user "heropup":

$$f(x) = \int_0^\infty f(x|\lambda) * g(\lambda) d\lambda$$

$$= \int_0^\infty \lambda e^{-\lambda x} \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} d\lambda$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \lambda^\alpha e^{-(\beta+x)\lambda} d\lambda$$

The integral now looks similar to the Gamma function, so we use the substitution $u = (\beta + x) * \lambda$. Note that the limits of the integral do not change since this is just a scalar multiple and both *x* and β are positive. Continuing from above:

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \left(\frac{u}{\beta + x}\right)^\alpha e^{-u} \frac{du}{\beta + x}$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\beta + x}\right)^{\alpha+1} \int_0^\infty u^\alpha e^{-u} du$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\beta + x}\right)^{\alpha+1} \Gamma(\alpha + 1)$$

Using the identity $\Gamma(\alpha + 1) = \alpha * \Gamma(\alpha)$ we find:

$$f(x) = \frac{\alpha\beta^\alpha}{(\beta + x)^{\alpha+1}}$$

$$= \frac{\alpha}{\beta}\left(1 + \frac{x}{\beta}\right)^{-(\alpha+1)}$$

This is the PDF of the Lomax distribution. It is essentially a standard Pareto distribution except the values are the amount in excess of the scale or threshold, which in this case is β. If the reader wishes to align these distributions with the Pareto types [4], then the standard Pareto refenced above is Type I, and the Lomax is a special case of Type II having $\mu = 0$.

3.2 Gamma-Pareto Mixture

We showed before that the Pareto is "Log-Exponential". Suppose we have a Pareto distributed random variable x with a fixed scale or threshold, but the Pareto shape or λ parameter is Gamma distributed. We have:

$$F(x) = 1 - \left(\frac{x}{\theta}\right)^{-\lambda} ; x > \theta$$

$$\mathrm{g}(\lambda; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$$

Using the substitution $y = \ln(x/\theta)$, from the previous section we know that y, unconditional on λ, is Lomax distributed with parameters α, β. Referencing Wikipedia for the CDF of the Lomax [3], we have:

$$F(x) = 1 - \left(1 + \frac{\ln\left(\frac{x}{\theta}\right)}{\beta}\right)^{-\alpha}$$

Note that the PDF would not just be plugging the substitution into the PDF of the Lomax for *y*. The PDF is the derivative of the CDF, and so must contain an *x* somewhere due to the logarithm term. For completion, taking the derivative we obtain:

$$f(x) = \alpha \left(1 + \frac{\ln\left(\frac{x}{\theta}\right)}{\beta}\right)^{-(\alpha+1)} * \frac{\theta}{\beta x}$$

4. MOMENTS AND LIMITED EXPECTED VALUES

The introduction mentions that the mean and all moments of the "Log-Lomax" distribution are infinite. We can see this in two ways. First, we recall that the Pareto distribution only has a finite mean for shape parameter $\lambda > 1$. For this mixture, we have a certain probability of all λ's from zero to infinity. Since there is a positive probability of having $\lambda \leq 1$, then the overall mean of the mixture must be at least that probability times infinity, hence it is infinite. The argument applies for all higher moments since the k^{th} moment of a Pareto only exists for $\lambda > k$.

The other way to see (which is actually the same reason mathematically) is that the PDF decays like:

$$f(x) \sim \frac{1}{x \ln(x)^{\alpha+1}}$$

Obviously if we multiply by x^k for integer $k > 1$, then this quantity actually increases, giving an infinite integral. But even for $k = 1$, i.e. the mean, the quantity $\ln(x)^{\alpha+1}$ grows more slowly than *x* for any positive α, hence we have for large enough *R*:

$$\int_R^\infty \frac{1}{\ln(x)^{\alpha+1}} dx \geq \int_R^\infty \frac{1}{x} dx = \infty$$

4.1 Limited Expected Value

Even though the mean and moments are infinite, the limited expected values must obviously be finite. Especially if we are interested in reinsurance pricing applications, typical quantities of interest would be the expected loss, average severity, and standard deviation of losses in a layer. Those can all be derived from the moments of the limited expected value of x.

Again suppose:

$$F(x) = 1 - \left(1 + \frac{\ln\left(\frac{x}{\theta}\right)}{\beta}\right)^{-\alpha}$$

We then calculate the limited expected value:

$$E((x \wedge L)^k) = \int_{\theta}^{\infty} \min(x, L)^k * f(x)dx$$

$$= \int_{\theta}^{L} x^k \frac{\alpha\theta}{\beta x}\left(1 + \frac{\ln\left(\frac{x}{\theta}\right)}{\beta}\right)^{-(\alpha+1)} dx + L^k * S(L)$$

For the second term we use the CDF formula from above:

$$L^k * S(L) = L^k\left(1 + \frac{\ln\left(\frac{L}{\theta}\right)}{\beta}\right)^{-\alpha}$$

The first term can be expressed as:

$$\frac{\alpha\theta}{\beta}\int_{\theta}^{L} x^{k-1}\left(1 + \frac{\ln\left(\frac{x}{\theta}\right)}{\beta}\right)^{-(\alpha+1)} dx$$

$$= \frac{\alpha\theta}{\beta}\int_{\theta}^{L} x^{k-1}\beta^{\alpha+1}\left(\beta + \ln\left(\frac{x}{\theta}\right)\right)^{-(\alpha+1)} dx$$

$$= \alpha\theta\beta^{\alpha}\int_{\theta}^{L} x^{k-1}\left(\beta + \ln\left(\frac{x}{\theta}\right)\right)^{-(\alpha+1)} dx$$

To simplify the integral, we make the substitution:

$$u = k(\beta + ln(x/\theta))$$

This gives us:

$$du = \frac{k\theta}{x}dx \Rightarrow dx = \frac{x}{k\theta}du$$

$$x = \theta e^{u/k-\beta}$$

$$x = \theta \Rightarrow u = \theta e^{\theta/k-\beta}$$

$$x = L \Rightarrow u = \theta e^{L/k-\beta}$$

Substituting those into the integral we get:

$$\alpha\theta\beta^{\alpha}\int_{\theta e^{\theta/k-\beta}}^{\theta e^{L/k-\beta}} x^{k-1}(u/k)^{-(\alpha+1)}\frac{x}{k\theta}du$$

$$= \frac{\alpha\beta^{\alpha}}{k^{\alpha+2}}\int_{\theta e^{\theta/k-\beta}}^{\theta e^{L/k-\beta}} x^{k}(u)^{-(\alpha+1)}du$$

$$= \alpha\beta^{\alpha}\int_{\theta e^{\theta/k-\beta}}^{\theta e^{L/k-\beta}} (\theta e^{u/k-\beta})^{k}(u)^{-(\alpha+1)}du$$

$$= \alpha\beta^{\alpha}\theta^{k}e^{-k\beta}\int_{\theta e^{\theta/k-\beta}}^{\theta e^{L/k-\beta}} e^{u}u^{-(\alpha+1)}du$$

Unfortunately, the integral part of the expression above does not have a closed form. It can be approximated numerically, or as we show below it can be expressed as a difference of values in the Incomplete Gamma Function (whose values themselves are numerically approximated).

Plugging the results back into the original expression for the limited expected value we obtain:

$$E((x \wedge L)^{k}) = \alpha\beta^{\alpha}\theta^{k}e^{-k\beta}\int_{\theta e^{\theta/k-\beta}}^{\theta e^{L/k-\beta}} e^{u}u^{-(\alpha+1)}du + L^{k}\left(1+\frac{\ln\left(\frac{L}{\theta}\right)}{\beta}\right)^{-\alpha}$$

4.2 Incomplete Gamma Function

The unsolvable integral from above is of the form:

$$\int_a^b e^x x^{-c}\, dx$$

This looks very similar to the Gamma function. One could imagine substituting $u = -x$ and getting:

$$\int_{-b}^{-a} e^{-u}(-u)^{-c}\, du$$

Suspending disbelief momentarily, let's assume all the integrals and quantities involved exist and are finite. If the lower incomplete gamma function, defined by:

$$\Gamma(a, x) = \int_x^\infty e^{-t} t^{a-1} dt$$

existed for negative values of *x*, then we could "simplify" the unsolvable integral above as:

$$\int_{-b}^{-a} e^{-u}(-u)^{-c}\, du = (-1)^{-c} \int_{-b}^{-a} e^{-u} u^{-c}\, du$$
$$= (-1)^{-c}(\Gamma(-c+1, -b) - \Gamma(-c+1, -a))$$

The two questions we need to ask are: is this true, and is this helpful? The problem with this being true is that we would need to know that $\Gamma(a, x)$ exists for negative *x*. Gautschi [3, pp. 3-4] references earlier works (that were unavailable to this author) on the incomplete gamma function which define:

$$\gamma^*(a, x) = \frac{x^{-a}}{\Gamma(a)} \gamma(a, x)$$

where:

$$\gamma(a, x) = \int_0^x e^{-t} t^{a-1} dt = \Gamma(a) - \Gamma(a, x)$$

and show that $\gamma^*$is real valued for real *a* and *x* and exists for *x*<0. It's well established that gamma functions can be extended to negative *a* values (excluding negative integers) using recurrence relations arising from integration by parts.

Putting these results together, we can see how the $(-1)^{-c}$ term of our expression will be cancelled out by the x^{-a} term in the definition of $\gamma^*(a, x)$ to show that the representation of the integral by the incomplete gamma function will exist and be real.

But is this useful? Unfortunately, no for a couple of reasons. First, not all tools, programs and packages return values for negative arguments of the incomplete gamma function. Sage (available through cocalc.org) and some online calculators do, but standard packages in R and Python do not. Secondly, the size of some of the terms becomes large enough that the accuracy of the program used becomes questionable. Even in a simple example tested, the $\alpha\beta^\alpha\theta^k e^{-k\beta}$ term became something on the order of 10^{70}, and this was after attempting to use multiple tools, some of which returned an overflow error. Should we really trust if a program tells us $1.234 \ldots * 10^{70} - 1.234 \ldots * 10^{70} = 42$?

However useful or not, the closed form formula for limited moments of the Pareto-Gamma mixture is given by the following:

$$E((x \wedge L)^k) = \alpha\beta^\alpha\theta^k e^{-k\beta}(-1)^{-(\alpha+1)}\left(\Gamma\left(-\alpha, -\theta e^{\frac{L}{k}-\beta}\right) - \Gamma\left(-\alpha, -\theta e^{\frac{\theta}{k}-\beta}\right)\right)$$

$$+L^k\left(1+\frac{\ln\left(\frac{L}{\theta}\right)}{\beta}\right)^{-\alpha}$$

Due to these constraints, the author recommends approximating quantities of interest by simulating random draws and taking a sample average. Take care however not to do this with quantities whose true value is infinite, e.g. the uncapped mean of the mixture.

5. CONCLUSION

The Pareto-Gamma mixture can arise in the context of Bayesian models for large losses. We have shown some properties of the distribution, and derived closed form formulas for some quantities of interest, although unfortunately some of them are not of practical use. Hopefully the availability of this information can aid and encourage the recognition of parameter risk in large loss modeling, or perhaps the discovery of a similar model matching empirical data, but with more tractable formulas for quantities of interest.

6. REFERENCES

[1] heropup, "compound of gamma and exponential distribution", StackExchange, https://math.stackexchange.com/questions/646852/compound-of-gamma-and-exponential-distribution/, 2014.

[2] Conjugate Prior. (January 14, 2021). In *Wikipedia*. https://en.wikipedia.org/wiki/Conjugate_prior

[3] Lomax Distribution. (May 21, 2020). In *Wikipedia*. https://en.wikipedia.org/wiki/Lomax_distribution

[4] Pareto Distribution. (January 13, 2021). In *Wikipedia*. https://en.wikipedia.org/wiki/Pareto_distribution

[5] Gautsche, Walter. "Incomplete Gamma Function Since Tricomi". In Tricomi's Ideas and Contemporary Applied Mathematics, Atti dei Convegni Lincei, n. 147, Accademia Nazionale dei Lincei, 1998. http://www.cs.purdue.edu/homes/wxg/Tricomi.ps

Abbreviations and notations

CDF, cumulative distribution function
PDF, probability distribution function

About the Author

Greg McNulty is Senior Vice President, Head of Americas Actuarial Pricing for SCOR Reinsurance and is based in Chicago. He has a BS in Mathematics from UCLA and an MS in Mathematics from the University of Michigan. He is a Fellow of the CAS. His prior works have been awarded the Ratemaking Prize and the Variance Prize.

The Single Parameter Pareto Revisited

Rajesh Sahasrabuddhe

January 21, 2021

Abstract

In his paper in the 1985 *Proceedings of the Casualty Actuarial Society*, Stephen W. Philbrick [Philbrick 1985] proposed the use of the Single Parameter Pareto distribution to model excess layer claim severity distributions. In this paper, we reintroduce the distribution. We provide guidance as to when it is appropriate to model claims using the Single Parameter Pareto and offer a new approach to estimate its parameter and identify the appropriate threshold.

Keywords Pareto, excess claims, severity models

1 Introduction

1.1 Overview

In one of his seminal works, Stephen W. Philbrick proposed an elegant solution to the complex problem of modeling claims amounts in excess layers. The elegance is the result of the following:

- He proposed a *single parameter* distribution to model claim layers where there may be limited amounts of data. Having only a single parameter maximizes the degrees of freedom of the model.
- The maximum likelihood estimator of the parameter has a straightforward derivation.
- Actuaries are able to easily calculate severity values of interest.

The solution involved the use a Pareto Type I distribution for claims above an excess threshold. Philbrick referred to this distribution as the "Single Parameter Pareto" (referred to as the SPP throughout this paper).

This paper expands on Philbrick's work in several ways by providing the following:

In the remainder of this section, we reintroduce the SPP.

In **Section 2**, we describe a test to determine when it is appropriate to model claims using the SPP. We extend this test to support the determination (rather than selection) of the excess claims threshold and determination of the Pareto parameter. We conclude this section with a recipe for analysis.

In **Section 3**, we review the Pareto parameter. We explore the relationships of the parameter to both the excess loss threshold and cost levels.

In **Section 4**, we discuss actuarial applications of the SPP and parameter values less than 1.

In **Section 5**, we provide concluding remarks.

In **Appendix A**, we present an errata to *Philbrick*.

In **Appendix B**, we provide a review of the various formulæ related to the SPP including supporting derivation.

In **Appendix C**, we provide the `R` code supporting the figures included in this paper.

1.2 Preliminaries

Philbrick's SPP is a special case of the Pareto Type I distribution which has the following cumulative distribution function:

$$F(x) = 1 - \left(\frac{k}{x}\right)^a \tag{1.2.1}$$

The scale parameter (k) in the Pareto Type I is not necessary in the SPP because the data are scaled. As such the SPP is equivalent to a Pareto Type I with scale parameter equal to 1 that is fit to transformed data. We then invert the transform to calculate quantities of interest.

1.3 The Single Parameter

Readers of *Philbrick* are often confused by the reference to the *single parameter*. After all, in Section III Philbrick initially presents the Pareto with two parameters (as we present in Equation 1.2.1), k and a, and then later adds that claims should be "normalized" by dividing by the "selected lower bound."

This presentation leaves many readers not understanding how the lower bound, k, "lost" its parameter status. *Philbrick* explains that this is because:

> Although there may be situations where this value must be estimated, in virtually all insurance applications this value will be selected in advance. (Philbrick, Section III)

We offer the alternative view that users of the SPP model should consider the process of normalizing the claims to be a transformation of the data rather than the application of a parameter. An analogous transformation occurs when we take (natural) logarithms. When we do that, we do not consider the base of the logarithm (e) to be a parameter. Similarly, we should not consider the lower bound to be a parameter.

To improve the clarity of this concept, we present Table 1 comparing the more traditional two parameter Pareto Type I and the SPP. We denote the raw claim amounts as observations of the random variable **C** and the normalized claim amounts as **Z**[1].

Distribution	Pareto Type I	SPP
Random Variable	Observed Claim Amount **C**	Normalized Claims Amount **Z**
Transformation	Not Applicable	$\mathbf{Z} = g(\mathbf{C})$ $g(\mathbf{C}) = \mathbf{C}/\text{lower bound}$
Parameters	$k > 0$ (scale); $a > 0$ (shape)	$q > 0$ (shape)
Domain	$[k, \infty]$	$[1, \infty]$
Density	$a\dfrac{k^a}{C^{a+1}}$	$qz^{-(q+1)}$

Table 1: The Pareto Type I and the SPP

The support of both the Pareto Type I and the SPP distribution are claims in excess of a threshold. The support for the former is all claim amounts greater than k. The support of the latter is all claim amounts greater than the lower bound which results in the normalized claim amounts greater than 1.

We can now work with model forms in the space of **Z** and then use g^{-1} to transform back into the space of **C**. We can also now present the density and distribution functions.

$$f(z) = qz^{-(q+1)} \tag{1.3.2}$$

$$F(z) = 1 - z^{-q} \tag{1.3.3}$$

In Appendix B.1, we provide the derivation of the distribution function (Equation (1.3.3)).

[1] Later in this paper, we advocate an approach that requires plotting data on an x, y coordinate system. We use **C** and **Z** to avoid confusion with that coordinate system.

2 When is it appropriate to use the SPP?

Philbrick introduced the SPP as a distribution to model excess claims. As such, the most common actuarial use of the SPP is in the modeling of claims in the tail of a distribution which is of interest when the tail is said to be "thick" or "heavy." Of course, the terms 'thick" or "heavy" have no formal definition.

The simplicity/elegance of the SPP has had the unintended consequence that the SPP is widely-used without an assessment to determine if and where the data follow a Pareto distribution. (We note that Philbrick did not include such an assessment.) We propose an "assessment approach" (as compared to Philbrick's "selection approach") in this paper. We begin by applying our proposed approach to normalized data. We then extend the concept to data that is not normalized. We recommend that latter approach for actuaries to use in fitting the Pareto model.

2.1 The Zipf Plot

Specifically, we note that Pareto-distributed data plot as a straight line on a **Zipf plot**[Cirillo 2013].

To construct a Zipf plot, we plot the (empirical) survival function on the y-axis and the data points on the (transformed) x-axis. Both axes are on a log[2] scale.

y-values From Equation (1.3.3), we recognize that the survival function is $1-(1-z^{-q}) = z^{-q}$ and the natural logarithm of the survival function is $-q \ln z$.

x-values We note that the x values of the Zipf plot are $\ln z$

We represent the linear relationship [3] of the y and x values as:

$$-q \log z \sim b_1 \log z \qquad (2.1.4)$$

Simplifying Equation 2.1.4, we have the straightforward observation that the coefficient of $\log z$ on the right hand side of the relationship (i.e., b_1) represents an estimator for the negative of the Pareto parameter (i.e., $-q$).

2.2 Zipf Plot Example Normalized Pseudo Data

In this section, we demonstrate the use of the Zipf plot using *normalized* pseudo data. We refer to this data as *normalized* consistent with the *Philbrick* definition raw values have been divided by the selected lower bound.

[2] All reference to logarithms that I include throughout this paper are natural logarithms.

[3] This notation indicates that the left side of the $\sim$ is a function of the right side of the $\sim$ without specifying any possible other terms of the relationship.

To generate that pseudo data, we assume that each of n observed points is located at the midpoint of evenly-spaced probability intervals[4] . That is, the empirical distribution and survival functions for the ith **ordered** point, $(z_{(i)}, i \in [1, n])$ are:

$$F(z_i) \quad = \frac{i - 0.5}{n} \tag{2.2.5}$$

$$S(z_i) \quad = \frac{n - i + 0.5}{n}. \tag{2.2.6}$$

From Equation (1.3.3), we recognize that associated normalized data points, z have values $S(z_i)^{(-1/q)}$. Also, importantly, we use $z_{(1)}$ to represent the first observed value. We later discuss the significance of $z_{(0)}$.

Then pseudo data points on the Zipf plot are:

$$(x_i, y_i) = (\ln z_{(i)}, \ln(\frac{n - i + 0.5}{n}))$$

where we now use $z_{(i)}$ to indicate the i-th order statistic of the data sample.

Then, we can use linear modeling tools to facilitate calculation of slope of line through the data points using ordinary least squares.

For the normalized data, using the following logic, we understand that the constant in the relationship (that is, the y-intercept) is, by definition, 0:

- Since the x-axis represents values of the $\log(z)$, we denote the minimum z-value as $z_{(0)}$.
- $F(z_{(0)}) = 0; S(z_{(0)}) = 1$
- The y-value at $z_{(0)}$ is $\ln(S(z_{(0)})$ which is equal to $\ln 1 = 0$.

Similarly we understand that $z_{(0)} = 1$ results in an x-value $= \ln(z_{(0)}) = 0$. We now recognize that the line fit to the point on the Zipf-plot passes through the origin.

We present Zipf plots using 100 pseudo-data points at q values of 0.5, 1.0, 1.5 and 2.0 in Figure 1 and include the least squares fitted line and the associated regression coefficient. In Appendix C.1, we present the `R` code used to generate Figure 1 which includes a function that can be used to generate Zipf-plots (adapted from [Cirillo 2013]).

2.3 The Lower Bound

Suppose however that the data were *not* normalized by a *selected* lower bound (we will denote that lower bound as **B** here). Then, we multiply each of the

[4] We emphasize that this is pseudo data meant only to support the reader's replication of the example. We acknowledge that this is not the only means through which one could generate pseudo data. In practice, we assume that the reader would be applying the recipe to observed data.

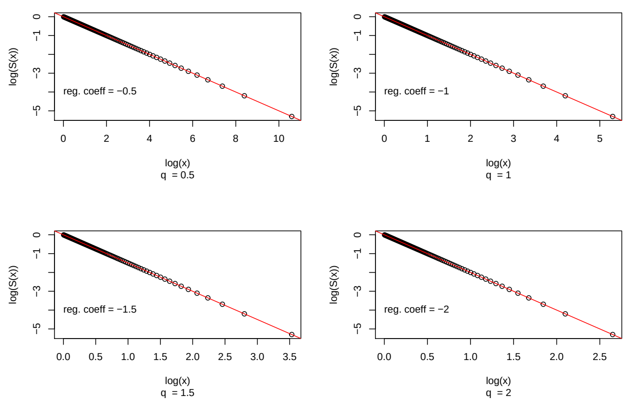


Figure 1: Zipf Plots

x-values in Figure 1 by $\mathbf{B}$, and the lns of the x-values would move rightward by $\ln B$. We can now see that we have returned to our original claim amounts $\mathbf{C}$. More importantly however, the constant in the linear relationship between $\ln \mathbf{c}$ and $\ln \mathbf{S}(\mathbf{z}) = \ln \mathbf{S}(\mathbf{c})$ is no longer 0 since the $c_{(0)}$ is now B rather than 0.

We can now use the linear relationship to solve for the x-value at which the y-value of the fitted line is equal to 0 (i.e., the x-intercept).

That is, we evaluate $\ln S(c) = \beta_0 + \beta_1 \log(c)$ at $c_{(0)}$.

$$\begin{aligned} \ln S(c_{(0)}) &= \beta_0 + \beta_1 \ln(x_{(0)}) \\ \ln 1 &= \beta_0 + \beta_1 \ln(c_{(0)}) \\ 0 &= \beta_0 + \beta_1 \ln(c_{(0)}) \\ c_{(0)} &= \exp(-\beta_0/\beta_1) \end{aligned}$$

Readers should recognize that we do not observe $c_{(0)}$; $c_{(1)}$ is our first observed value. The linear model provides statistical support for the threshold which in *Philbrick* was *selected*

2.4 Analysis Recipe

The elegance and stated purpose (modeling of excess layers) of the SPP invite its use without an evaluation of the appropriateness of the model. Key findings of this research paper are as follows:

Identify if and where data are Pareto-distributed Actuaries should use Zipf plots to understand first which data regions are indeed Pareto distributed. That is, the actuary should identify the range of data that exhibits a linear pattern on a Zipf plot. Statistical evaluation of linearity is outside the scope of this paper. However, we would suggest the following initial tests. (Note that below, we recommend a reevaluation of this assumption)

- a plot of the residuals of linear model as a function of fitted values
- a visual analysis

If the data do not appear to be linear on a Zipf plot, a Pareto model should not be selected.

If the data are (reasonably) linear in a certain region, the actuary, should then discard all data points outside the region of linearity.

Determine threshold and Pareto-parameter The actuary should then use the results of a linear model fit through the the remaining points (and only those points) of the Zipf-plot to parameterize that model. (This Zipf plot will differ from the initial plot as the empirical survival function will be calculated using only the retained data.)

- The negative of the covariate of $\ln c$ should be used as the estimator for the q parameter. (We explore advantages of this estimator to the maximum likelihood estimator in the next section.)
- the ratio of the negative of the intercept of the linear model and the covariate of $\ln c$ represents the value at which the data begin to be Pareto distributed.
- We can (and should) use tools (e.g., autocorrelation of residuals) that we use to statistically evaluate linear models to determine whether the data are indeed Pareto distributed (i.e., linear).
- The linear model output includes the standard error of the covariate (i.e., parameter uncertainty).

Application The actuary can then return to the presentation in Philbrick for formulæ relevant for modeling.

Actuarial judgment Actuaries should continue to apply judgment where appropriate throughout the process including but not limited to, assigning predictive value to underlying data points and in interpreting and using the modeling results.

Application of actuarial judgment is particularly important in addressing the practical issues that we discuss in Section 2.4.1.

2.4.1 Practical Issues

In executing this recipe in practice, there is one primary issue that we have encountered. That is, data are never perfectly Pareto-distributed as are the data in Figure 1. Specifically data will generally display some level of non-linearity. That is, data will typically display a concave down or concave up pattern.

Concave Downward Data that is concave downward will yield an indicated threshold that is greater than the initially selected threshold.

Concave Upward Data that is concave upward will produce an indicated threshold that is less that the initially selected threshold.

We have offer the following options for responding to this issue.

Re-select and Refit We could reexamine the data and identify a new value above which data are Pareto-distributed. We could then refit the linear model. This would be an iterative process.

Exclude individual data We could also exclude (inappropriately) influential points in fitting the model.

Accept the model In our experience, in many cases that data exhibit the patterns described, the indicated Pareto-parameter only changes slightly as we apply one of the two prior approached. So, we could retain the Pareto-parameter from the model but use judgment that considers both the model and the data in selecting the threshold value.

Actuaries should use professional judgment in the selection of which of these options to use.

3 The Pareto Parameter

Pseudo data plotted on Zipf plot also provides an important tool to help us understand the relationship between the data and the Pareto model. In this section, we use those tools to understand the following:

- Why the linear model produces a more robust estimator for the Pareto parameter.
- The effect of trend on the parameter and the lower bound.

3.1 Parameter Estimation

We now compare the approach that we present to Philbrick's approach to estimating the Pareto parameter.

- The maximum likelihood estimator (MLE) presented in *Philbrick*is:

$$q = \frac{n}{\sum \ln x}. \tag{3.1.7}$$

  We should recognize that the MLE is simply the reciprocal of the mean of $\ln x$.

- Option 2 is to use the coefficient of the linear model described in Section 2. For convenience, we will refer to this estimator as the CLM (coefficient of linear model).

We evaluate these alternatives considering the practical issues of missing data related to the estimation of the parameter. Specifically, we should understand that our observations may not be a representative sample of the claims that would be generated by a phenomenon that produces Pareto-distributed data.

To understand the effect on the parameter, we first consider situations where we only observed certain data points but that all possible data points are perfectly Pareto distributed. (That is, there is no process variance in the underlying data generation.) In Figure 2, we present an example where there are potentially 20 observed points (our population) and between 5 and 15 points are *not* observed in the samples (our samples). We generate all possible combinations under these conditions. In the upper panel, we plot the mean indicated parameter using MLE and CLM; in the lower panel, we plot the standard deviation of the indicated parameters.

We note that neither approach perfectly reproduces the underlying parameter. However, we did note that for many different values of the true parameter, the CLM resulted in an estimate closer to the actual value. We performed this analysis through simulation and present the underlying `R` code in Appendix C. Intuitively however we can recognize that a model (as is the basis for the CLM)

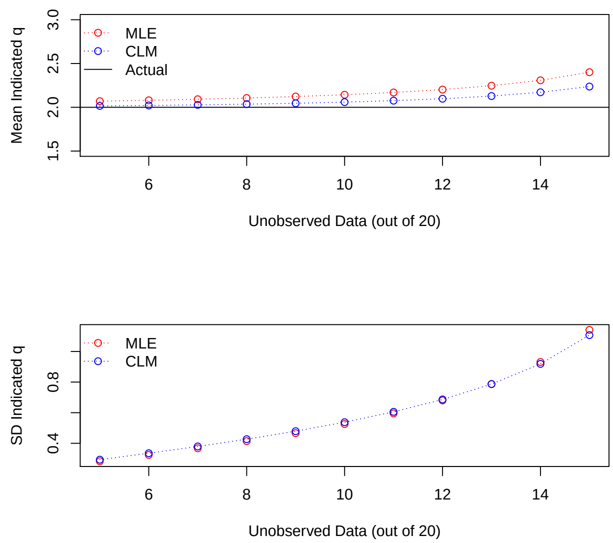


Figure 2: Comparison of Parameter Estimation Models

will help to extract signal from the data whereas an average (as this the basis for the MLE) effectively does not distinguish between noise and signal.

3.2 Trend

Philbrick espouses that the Pareto parameter should not be adjusted for claims inflation. That is, he argues that claims inflation results in frequency trend as more claim enter the "Pareto layer" but that there is no change to the SPP severity model.

With our recommended *assessment approach*, this is no longer intuitive or desirable. That is, Philbrick is implying that claims "become" Pareto-distributed once they trend to values above the threshold but that data is not Pareto-distributed below the threshold. In our approach, the use of the Zipf-plots maximizes the data used in our modeling. We identify the threshold above which all claims (which we presume have been appropriately adjusted to a common cost level) are Pareto-distributed. The Philbrick approach implies different levels of trend on either side of the threshold and that the trends acts in a way as to "push" claims over the threshold so as to preserve the Pareto parameter. While there is certainly a possibility that all these conditions are met, we view the simultaneous existence of all conditions as unlikely.

More specifically, we view "cost-leveling" as a separate modeling choice for the actuary. That modeling is outside the scope of this paper. Using the linear model-based approach to determining the Pareto parameter *and* the threshold, and assuming that each claim is subject to the same rate of trend, each claim in the Pareto-distributed portion of the data would move right (inflationary trend) or left (deflationary trend) by an amount equal to the ln of the trend adjustment. The x-intercept (that is the threshold value) would similarly shift and the y-coordinates would not change. Similarly the covariate would not change.

4 Actuarial Application

4.1 Parameter Values

To provide context to the Pareto parameter, we first review the application of the SPP.

We can use Equation (4.1.8), to calculate the limited expected value through b as presented in Appendix B.3 as:[5]

$$\mathbb{E}[X; b] = \frac{q - b^{1-q}}{q - 1} \tag{4.1.8}$$

In that derivation, there is no restriction that $q > 1$ as exists in the determination of the unlimited mean presented in Equation (B.2.14). That is, we recognize that, although the limited expected value is undefined, expected values are defined when we have an upper limit (such as a policy limit). Moreover, In Section IV., Philbrick indicates that:

> ... but most actual data suggests that the tail of the Pareto is still somewhat too thick at extremely high loss amounts. In other words, the theoretical density at high loss amounts is larger than empirical experience tends to indicate. Rather than discard the Pareto, it is easier to postulate that the distribution is censored or truncated at some high, but finite, value. As we have seen earlier, any upper limitation (either censorship point or truncation point) will produce formulæ for the mean claim size that are finite for all possible values of q.

As such, users of the SPP need not "fear" q values less than 1 for most insurance applications.

4.2 Claim Costs by Layer

Estimating claims for an excess policy is, of course, likely the most common use of the SPP. This was also a focus of Section III of Philbrick. For the expected claim amount for the layer between AP and L, we have:

[5] Philbrick used b to refer to both the "lower bound" and the policy limit. We will not do that in this paper primarily for clarity as using a variable to represent the lower bound implied at least the possibility that the lower bound was a parameter. Conveniently, it also allows us to use the traditional policy notation as attaching at AP through limit L with the resulting layer width equal to $L - AP$.

$$
\begin{aligned}
\mathbb{E}[X; AP, L] &= \frac{q - L^{1-q}}{q-1} - \frac{q - AP^{1-q}}{q-1} \\
&= \frac{AP^{1-q} - L^{1-q}}{q-1} \qquad (4.2.9)
\end{aligned}
$$

4.3 Policy Claims Estimate

The purpose of the Philbrick calculation was likely to demonstrate that the average claim size in the layer between AP and L was equal to the expected value of claims limited to L/AP net of the lower bound but multiplied by AP. The latter is calculated as Equation (4.1.8) -1 which simplifies to:

$$
\frac{1 - b^{1-q}}{q-1} \times AP \qquad (4.3.10)
$$

We can demonstrate that using Equation (4.2.9) and the survival function as follows:

$$
\begin{aligned}
\frac{\frac{AP^{1-q} - L^{1-q}}{q-1}}{S(AP)} &= \frac{\frac{AP^{1-q} - L^{1-q}}{q-1}}{AP^{-q}} \\
&= \frac{1}{q-1} \times \frac{AP}{AP} \times \frac{AP^{1-q} - L^{1-q}}{AP^{-q}} \\
&= \frac{AP}{q-1} \times \frac{AP^{1-q} - L^{1-q}}{AP^{1-q}} \\
&= \frac{AP}{q-1} \times \left(1 - \left(\frac{L}{AP}\right)^{1-q}\right) \\
&= \frac{1 - (L/AP)^{1-q}}{q-1} \times AP \qquad (4.3.11)
\end{aligned}
$$

As mentioned, the most common actuarial application of the SPP is to estimate the number of claims, their average value and the resulting aggregate claim amount to a policy. We summarize those formulæ for N ground-up claims in Table 2.

Number of Claims	$S(AP) = N \times AP^{-q}$
Average Value of Individual Claims	$\frac{1-(L/AP)^{1-q}}{q-1} \times AP$
Aggregate Claim Amount	$N \times \frac{AP^{1-q} - L^{1-q}}{q-1}$

Table 2: Policy Analysis

5 Concluding Remarks

Our goal with this paper was to provide additional guidance in deploying Philbrick's elegant solution to a complex problem. Our guidance supplements Philbrick with data visualization and model fitting that we expect would produce more robust solutions to the application of the Single Parameter Pareto in modeling excess claim layers.

Appendices

A Errata

In reviewing Philbrick, we noted two typographical errors and one calculation error. These are discussed below.

A.1 Philbrick Errata #1

In the application of formula (4.1.8), we should understand that there is a minor typographical error in Philbrick. The second paragraph following Equation (6) appears on Page 56 and includes the following:

$$\begin{gathered} b = 20 \times (500,000/25,000) \\ \text{which should be} \\ b = 20 = 500,000/25,000 \end{gathered}$$

A.2 Philbrick Errata #2

Starting at the bottom of Page 58 and extending to Page 59, Philbrick presents an example with a q parameter of 1.5 and expected claim count of 7 that results in the following (where $S(x)$ represents the survival function):

$$\begin{aligned} F(4) &= 1 - 4^{-1.5} \\ F(4) &= 7/8 \\ S(4) &= 1 - F(4) = 1/8 \end{aligned}$$

$$\begin{aligned} \mathbb{E}[n] &= 7 \\ \mathbb{E}[n; x > 4] &= 7 \times S(4) = 7/8 \end{aligned} \tag{A.2.12}$$

(It is unfortunate that, in this example both $\mathbb{E}[n; x > 4]$ and $S(4)$ both equal 7/8.)

$$\begin{aligned} \mathbb{E}[X] &= \frac{1.5}{1.5 - 1} \\ \mathbb{E}[X] &= 3 \end{aligned}$$

$$
\begin{aligned}
\mathbb{E}[X;4] &= \frac{1.5 - 4^{1-1.5}}{1.5-1} \\
\mathbb{E}[X;4] &= \frac{1.5 - 4^{-0.5}}{0.5} \\
\mathbb{E}[X;4] &= \frac{1.5 - .5}{0.5} \\
\mathbb{E}[X;4] &= 2
\end{aligned}
$$

The average severity of claims in the layer is $(\mathbb{E}[X] - \mathbb{E}[X;4])/S(4) = 8$. Using the frequency calculated in Equation (A.2.12), we estimate claims in the layer to be $8 \times 7/8 = 7$ which agrees with Philbrick's calculation.

The error occurs when the example is extended to calculate claims in the layer from $AP = 3$ to $L = 7.5$. Using the approach above, we have the following:

$$
\begin{aligned}
\mathbb{E}[X;3] &= 1.845299 \\
\mathbb{E}[X;7.5] &= 2.269703 \\
F(3) &= 0.8075499 \\
S(3) &= 0.1924501.
\end{aligned}
$$

We have average claim amounts in the layer at

$$
\frac{\mathbb{E}[X;7.5] - \mathbb{E}[X;3]}{1 - F(3)} = 2.205267
$$

which agrees with Philbrick's calculation of "net average claim size" on Page 59. However, the corresponding frequency should be $7 \times S(3) = 1.347151$ and resulting expected claims in the layer of 2.970827. The purpose of the $F(87,500/75,000) = F(2.5)$ term in the frequency calculation is not entirely clear to this author.

A.3 Philbrick Errata #3

Equation (11) indicates that "nth moment of the Pareto distribution with no upper limit is" $\frac{q}{q+n}$. Then, in Equation (12) the second moment is represented in the calculation of variance by $\frac{q}{q-n}$ and of course we have the calculation of mean (first moment, $n = 1$) as $\frac{q}{q-1}$. We can see the error in Equation (11).

B Derivation of Forumulæ

B.1 SPP Cumulative Distribution Function

$$
\begin{aligned}
F(x) &= \int_1^x f(x)\,\mathrm{d}x \\
&= \int_1^x qx^{-(q+1)}\,\mathrm{d}x \\
&= q\int_1^x x^{-(q+1)}\,\mathrm{d}x \\
&= q\frac{1}{-(q+1)+1}x^{-q}\Big|_1^x \\
&= q\frac{1}{-q}x^{-q}\Big|_1^x \\
&= -x^{-q}\Big|_1^x \\
&= -x^{-q} - (-1^{-q}) \\
F(x) &= 1 - x^{-q} \qquad \text{(B.1.13)}
\end{aligned}
$$

B.2 Expected Values

$$
\begin{aligned}
\mathbb{E}[X] &= \int_1^\infty xf(x)\,\mathrm{d}x \\
&= \int_1^\infty xqx^{-(q+1)}\,\mathrm{d}x \\
&= q\int_1^\infty x^{-q}\,\mathrm{d}x \\
&= q\frac{1}{-q+1}x^{-q+1}\Big|_1^\infty \\
&= \frac{q}{1-q}x^{-q+1}\Big|_1^\infty \\
\mathbb{E}[X] &= \frac{q}{1-q}\frac{1}{x^{q-1}}\Big|_1^\infty \qquad \text{(B.2.14)}
\end{aligned}
$$

We can see that for $x = 1$ (the lower limit of integration) equation (B.2.14) evaluates to $\frac{q}{1-q}$. However for $x = \infty$ (the upper limit of integration), we have the following[6]:

[6] In the limit as $x \to \infty$, the expression evaluates to $-\frac{q}{q+1}$. However evaluated at ∞, the expression is undefined.

$$\frac{q}{1-q}\frac{1}{x^{q-1}} = \begin{cases} 0, & \text{if } q > 1 \\ \text{undefined}, & \text{if } q = 1 \\ \infty, & \text{if } q < 1 \end{cases}$$

and therefore we have:

$$\mathbb{E}[X] = \begin{cases} 0 - \dfrac{q}{1-q}, & \text{if } q > 1 \\ \text{undefined}, & \text{if } q = 1 \\ \infty, & \text{if } q < 1 \end{cases}$$

or more simply:

$$\mathbb{E}[X] = \begin{cases} \dfrac{q}{q-1}, & \text{if } q > 1 \\ \text{undefined}, & \text{if } q \leq 1 \end{cases} \tag{B.2.15}$$

B.3 SPP Limited Expected Value

The limited expected value is calculated as:

$$
\begin{aligned}
\mathbb{E}[X;b] &= \int_1^b xf(x)\,\mathrm{d}x + b(1-F(b)) \\
&= \int_1^b xqx^{-(q+1)}\,\mathrm{d}x + b(1-F(b)) \\
&= q\int_1^b x^{-q}\,\mathrm{d}x + b(1-F(b)) \\
&= q\frac{1}{-q+1}x^{-q+1}\Big|_1^b + b(1-F(b)) \\
&= \frac{q}{1-q}\frac{1}{x^{q-1}}\Big|_1^b + b(1-F(b)) \\
&= \frac{q}{1-q}\left[\frac{1}{b^{q-1}} - \frac{1}{1^{q-1}}\right] + b\left[1-(1-b^{-q})\right] \\
&= \frac{q}{1-q}\left[\frac{1}{b^{q-1}} - 1\right] + b\left[b^{-q}\right] \\
&= \frac{q}{q-1}\left[1 - \frac{1}{b^{q-1}}\right] + b^{1-q} \\
&= \frac{q}{q-1}\left[1 - b^{1-q}\right] + b^{1-q} \\
&= \frac{1}{q-1}\left[q - qb^{1-q} + (q-1)b^{1-q}\right] \\
&= \frac{1}{q-1}\left[q - b^{1-q}\right] \\
&= \frac{q-b^{1-q}}{q-1} \qquad \text{(B.3.16)}
\end{aligned}
$$

B.4 Maximum Likelihood Estimator for Parameter

The negative log-likelihood (NLL) function given data $D = x_1 \ldots x_n$ is defined as:

$$
\begin{aligned}
L(q) &= \prod_{i=1}^{n} f x_i \\
\text{NLL} &= -\sum_{i=1}^{n} \ln(f x_i) \\
\text{NLL} &= -\sum_{i=1}^{n} \ln(q x_i^{-(q+1)}) \\
\text{NLL} &= -\sum_{i=1}^{n} \left[\ln q + \ln x_i^{-(q+1)}\right] \\
\text{NLL} &= -\sum_{i=1}^{n} \left[\ln q - (q+1)\ln x_i\right] \\
\text{NLL} &= -n\ln q + \sum_{i=1}^{n} (q+1)\ln x_i \\
\text{NLL} &= -n\ln q + (q+1)\sum_{i=1}^{n} \ln x_i
\end{aligned}
$$

We can calculate the MLE of q by taking partial derivatives and setting equal to 0.

$$
\begin{aligned}
0 &= \frac{\partial}{\partial q}\left[-n\ln q + (q+1)\sum_{i=1}^{n} \ln x_i\right] \\
0 &= -n\frac{1}{q} + \sum_{i=1}^{n} \ln x \\
\sum_{i=1}^{n} \ln x &= \frac{n}{q} \\
q &= \frac{n}{\sum_{i=1}^{n} \ln x}
\end{aligned}
$$

C R Code

We present the R code used to generate Figure 1 and Figure 2 below.

C.1 R Code for Figure 1

```
zipfplot <- function(data) {
  data <- x_values
  data <- sort(as.numeric(data)) #sorting data
  y <- 1 - ppoints(data) # computing 1-F(x)

  plot(x = data, y = y, log = 'xy', xlab = 'x on log scale',
    ylab = '1-F(x) on log scale')
}

n_points <- 100

y_vals <- ( n_points - (n_points:1) + 2 / n_points ) /
  n_points

par(mfrow = c(2,2))

x_vals <- y_vals

lapply(X = c(0.5, 1, 1.5, 2), FUN = function(q){

  #q <- 2

  plot(x = log(x_vals ^ (-1/q)), y = log(y_vals),
    xlab = 'log(x)',
    ylab = 'log(S(x))', sub = paste0('q  = ', q))

  fit <- lm(log(y_vals) ~ log(x_vals ^ (-1/q)) + 0)

  abline(fit, col = 'red')

  text(x = 0, y = -4, labels = paste0('reg. coeff = ',
    round(fit$coefficient, 1)),
    adj = 0)
})
```

C.2 R Code for Figure 2

```
# This may take a reasonably long time to run!
missing_pts <- 5:15

n_points <- 20

scale <- 200000

q <- 2

y_vals <- (n_points - (n_points:0)) / n_points
y_vals <- (y_vals[1:(length(y_vals) - 1)] +
  y_vals[2:length(y_vals)]) / 2

x_values <- (1 - y_vals) ^ (-1 / q)
rm(y_vals)

mle <- lapply(X = missing_pts, FUN = function(missing) {
  no_sampled_points <- n_points - missing

  combn(x = x_values, m = no_sampled_points,
    FUN = function(sampled_points) {
      no_sampled_points / sum(log(sampled_points))
    }
  )
})


clm <- lapply(X = missing_pts,   FUN = function(missing){

  no_sampled_points <- n_points - missing

  combn(x = x_values, m = no_sampled_points,
    FUN = function(sampled_points) {
      sampled_points <- sampled_points[order(sampled_points)]

      y_vals <- (no_sampled_points - (no_sampled_points:0) ) /
        no_sampled_points
      y_vals <- (y_vals[1:(length(y_vals) - 1)] +
        y_vals[2:length(y_vals)]) / 2
      y_vals <- 1 - y_vals

      -lm(log(y_vals) ~  log(sampled_points) + 0)$coefficient
    }
```

```
  )
})


clm_mean <- sapply(clm, mean, simplify = TRUE)
mle_mean <- sapply(mle, mean, simplify = TRUE)
clm_sd <- sapply(clm, sd, simplify = TRUE)
mle_sd <- sapply(mle, sd, simplify = TRUE)
save(clm, mle, clm_mean, clm_sd, mle_mean, mle_sd,
  file = './param_test.RData')

par(mfrow = c(2, 1))
plot(x = missing_pts, y = sapply(mle, mean, simplify = TRUE),
  ylab = 'Mean Indicated q',
  xlab = 'Number of Unobserved Data Points (out of 20)',
  type = 'n', ylim = c(1.5, 3))
abline(h = 2, lty = 'solid')
points(x = missing_pts, y = mle_mean, col = 'red')
lines(x = missing_pts, y = mle_mean, col = 'red',
  lty = 'dotted')
points(x = missing_pts, y = clm_mean, col = 'blue')
lines(x = missing_pts, y = clm_mean, col = 'blue',
  lty = 'dotted')
legend('topleft', legend = c('MLE', 'CLM', 'Actual'),
  lty = c('dotted', 'dotted', 'solid'), pch = c(1, 1, NA),
  col = c('red','blue', 'black'), bty = 'n')

plot(x = missing_pts, y = sapply(mle, sd, simplify = TRUE),
  ylab = 'SD Indicated q',
  xlab = 'Number of Unobserved Data Points (out of 20)',
  type = 'n')
points(x = missing_pts, y = sapply(mle, sd, simplify = TRUE),
  col = 'red')
lines(x = missing_pts, y = sapply(mle, mean, simplify = TRUE),
  col = 'red', lty = 'dotted')
points(x = missing_pts, y = sapply(clm, sd, simplify = TRUE),
  col = 'blue')
lines(x = missing_pts, y = sapply(clm, sd, simplify = TRUE),
  col = 'blue', lty = 'dotted')
legend('topleft', legend = c('MLE', 'CLM'),
  lty = c('dotted', 'dotted'),
  pch = c(1, 1), col = c('red','blue'), bty = 'n')
```

References

[1] Pasquale Cirillo. "Are Your Data Really Pareto Distributed?" In: *Physica A: Statistical Mechanics and its Applications* 392 (23 2013), pp. 5947–5962. URL: `https://arxiv.org/abs/1306.0100`.

[2] Stephen W. Philbrick. "A Practical Guide to the Single Parameter Pareto Distribution". In: *Proceedings of the Casualty Actuarial Society* LXXII (1985), pp. 44–84.

IFRS17 Measurement of Property & Casualty Insurance Contracts

Brett Ward, BEc, FIAA

Abstract: IFRS 17 introduces the concept of a risk adjustment that compensates insurers for the uncertainty about the amount and timing of the cash flows that arise from non-financial risks. The method for its calculation is not prescribed and several options are emerging, including value at risk and cost of capital. This paper recalls (Myers & Cohn, 1981) to provide a cost of capital approach that has desirable characteristics including relative ease of implementation, risk adjustment margins that are fully diversified and additive from granular levels, alignment with pricing bases and recognition of a uniform return on allocated capital.

1. INTRODUCTION

1.1. New accounting standard

IFRS17 Insurance Contracts will replace *IFRS4* for annual reporting periods from 1 January 2023 in setting out principles for the recognition, measurement, presentation and disclosure of insurance contracts. A further objective is to achieve greater consistency in financial reporting for life, health and property & casualty insurers, as well as greater consistency with other industries.

IFRS17 Paragraph 32 requires that measurement using the General Measurement Model (**GMM**) of insurance contracts on initial recognition be the total of:

(a) the fulfilment cash flows, which comprise:
   (i) estimates of future cash flows;
   (ii) an adjustment to reflect the time value of money and the financial risks related to the future cash flows; and
   (iii) a risk adjustment for non-financial risk.

(b) the contractual service margin.

Remeasurement at reporting periods after contract inception follows similar principles. IFRS17 also includes a simplified Premium Allocation Approach (**PAA**) for the measurement of insurance contracts where the coverage period is one year or less. It is expected that many insurers will adopt the PAA for their property & casualty insurance contracts.

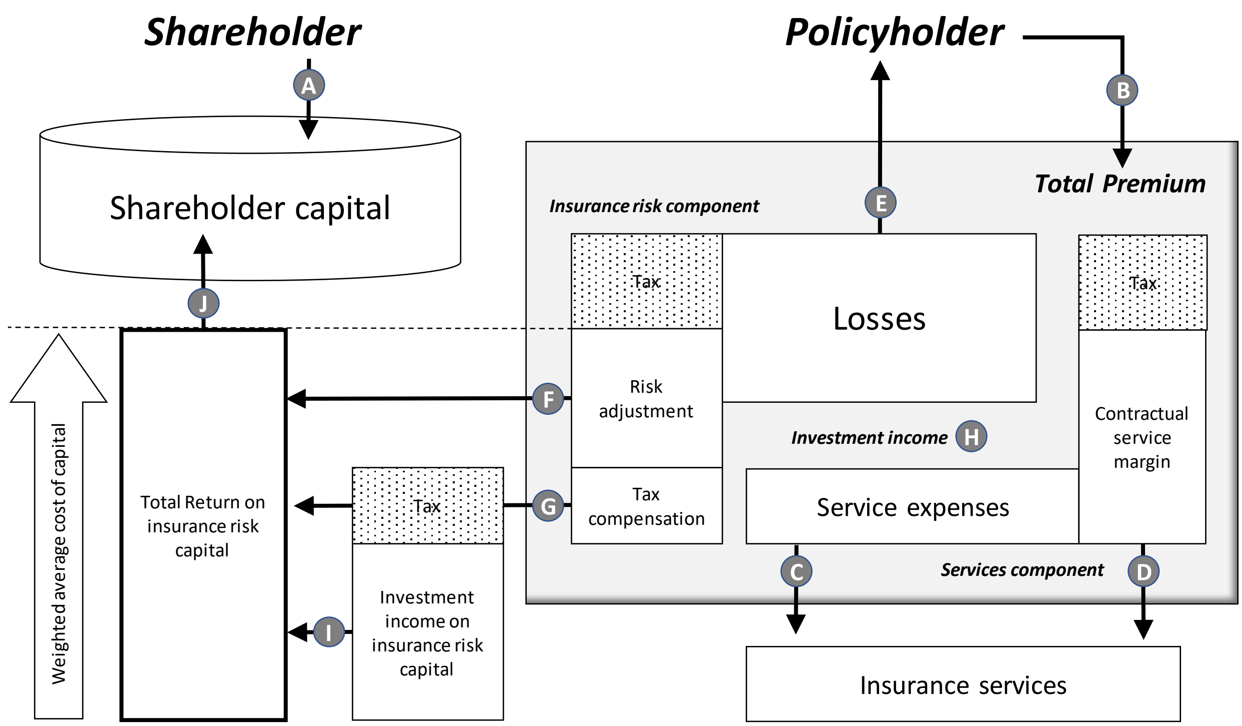
A key change that the IFRS17 measurement approach represents to property & casualty insurers is the use of the risk adjustment (**RA**) and the contractual service margin (**CSM**) to determine the recognition of profit.

This paper describes an approach to measurement of a property & casualty insurance contract that satisfies the requirements of IFRS17.

1.2. Insurance contract cashflows

Before exploring measurement of insurance contracts under IFRS17, it is useful to define the key cashflows that arise and understand their nature and purpose. The following diagram explores the cashflows associated with a typical property & casualty insurance contract between an insurer and policyholder. Also included are the cashflows involving the shareholder who supports the insurance contract with capital and receives a return. Each cashflow, marked with a letter, is discussed in some detail below.

Diagram 1: Property & Casualty Insurance Contract Cashflows



A. As cashflows are uncertain in their timing and ultimate amount, capital is needed to act as a buffer against adverse cashflow movements and enable the insurer to continue to fulfil its insurance contract liabilities and provide services. Cashflow **A** represents an injection into the insurer of capital by shareholders that will be allocated to support each insurance contract. For the purpose of this paper, there are two elements of capital allocated, that for insurance risk and that for the operational risk of the services the insurer provides.

B. An insurer and a policyholder enter into an insurance contract where the policyholder will be paid losses they may incur from insured events. Policyholders also receive a range of services associated with their contract that include distribution, product design & underwriting, claim handling, supply chain access and corporate services that support the operation of the insurer. The policyholder will pay a total premium, cashflow **B**, to the insurer in exchange for the insurance contract.

C. Services within the contract may be provided by the insurer or a range of external suppliers. The expense of services may be incurred prior to inception of the contract (such as product design), during the process of inception (such as distribution and underwriting) or after inception (such as claim handling and supply chain access). The service providers are compensated for their expenses by cashflow **C**.

D. Services may also receive a contractual service or profit margin, cashflow **D**, should they have provided value or utility to the policyholder in excess of the service expense that the policyholder was prepared to pay for. This profit ought to be released to the service proportional to service delivery. Before the profit is released though, a portion is paid as tax. This profit also includes the return on operational risk capital allocated to the insurance contract.

The value at inception of service expenses and contractual service margin, cashflows C and D, are referred to in this paper as the *Services* component of the total premium.

E. Policyholders experiencing a loss receive a payment, cashflow **E.** Although insured events will typically need to arise during the coverage time boundaries of the insurance contract, losses may be paid some time afterwards.

F. Policyholders need to reward, being cashflow **F**, the shareholder for the risk associated with the capital allocated to the contract for insurance risk. This is referred to as a risk adjustment. A portion of the risk adjustment is paid as tax.

G. Policyholders also need to compensate, cashflow **G**, the shareholder for tax that the insurer incurs on the investment income earned by the insurance risk capital. Shareholders would otherwise experience an extra layer of tax compared to investing directly in those same market assets. To make it at least a neutral proposition (before considering risk) to inject capital into an insurer compared to investing directly in the market, policyholders compensate the shareholder for this extra layer of tax through additional premium. As this additional premium is also taxed, the policyholder needs to provide a grossed-up compensation such that when tax is applied, the residual exactly compensates the tax on investment income on the insurance risk capital.

To illustrate the tax compensation component further, assume that a shareholder could invest $1,000 directly in an asset that provides an investment income of $70. Instead, this shareholder has decided to provide their funds to an insurer who now acquires the same asset in addition to writing insurance contracts. The investment income of $70 will be taxed in the hands of the insurer. Assuming the corporate tax rate is 30%, then this means the insurer pays $21 in tax leaving $49 to pass on to the shareholder. This is not acceptable to the shareholder who expects that the investment income on their $1,000 would be at least $70 plus a reward for risk from the insurance contracts. The policyholder is therefore required to compensate the shareholder for the $21 paid as tax through paying an additional premium. As premium is taxed, then the policyholder pays an additional premium amount equal to a grossed-up amount of $21 ÷ (1—30%) = $30.

Hence the $30 of additional premium is taxed at 30%, leaving $21 to add to the after-tax investment income of $49 thus providing a total of $70 for the shareholder – the same return as investing directly in the asset.

The value at inception of losses, the risk adjustment and tax compensation, cashflows E, F and G, are referred to in this paper as the *Insurance risk* component of the total premium.

H. The insurance contract cashflows C, D, E, F & G are not all paid at the point of inception. The total premium less any cashflows that occur at inception will be invested and produce investment income as cashflow **H** over time. This investment income has been allowed for in the present value of the insurance contract cashflows and is hence absorbed in the fulfilment of those cashflows.

I. Insurance risk capital is invested to produce an investment income as cashflow **I**. This investment income will be subject to tax, which is compensated for by the policyholder through cashflow G.

J. The final cashflow **J** represents the total return on insurance risk capital and is the sum of cashflows F, G and I – less the tax that is payable on all three of these cashflows. This total return has an expectation at inception of the contract of being equal to the weighted average cost of capital for the shareholder. If it were a lower return, then shareholders would look for other opportunities that did meet their cost of capital expectations. A higher return may be unfair on the policyholder who would be paying a higher premium than would be required to attract enough capital to support the risk in the contract.

1.3. The impact of tax on insurance contracts and insurance risk capital

The impact of tax on an insurance contract and capital allocated is significant. Tax scales down all cashflows in the same way a reinsurance quota share contract in the proportion of the corporate tax rate would take a share of all cashflows. Tax hence scales down the capital required to support an insurance contract compared to a tax-free environment.

To illustrate this further, assume an insurer underwrites two identical insurance contracts where one is in a tax-free environment, but the other is in an environment with a corporate tax rate of 30%. Suppose the insurance risk capital needed to support the contract in the tax-free environment is $1,000, for one year, and the risk adjustment component of the premium is $50. The shareholder receiving this risk adjustment therefore earns a 5% return on capital for assuming the risk under the insurance contract. The shareholder will also earn investment income through investment of the capital, but this is omitted in this simplified example. In the environment with the 30% corporate tax rate, all the insurance contract cashflows are scaled down 30%. It is as though the insurance contract is (1-30%) = 70% of the size of the contract in the tax-free environment. The shareholder is therefore required to inject only 70% of the capital in the taxed environment compared to the tax-free environment, or 70% × $1,000 = $700. The risk adjustment component of $50 is also taxed at 30% and hence the shareholder receives $35 that remains a 5% return on the $700 of capital injected. The return for the risk assumed is the same in the tax free and taxed environments, but the capital required is lower in the taxed environment to the extent of the corporate tax rate. The policyholder pays the same risk adjustment of $50 regardless of the level of the corporate tax rate.

1.4. Recognition of profit

A role of measurement under a financial reporting accounting standard is to recognise profit in proportion to the provision of insurance risk capital and delivery of services. Three elements to this profit have been identified:

- *Risk adjustment*: a cashflow from the policyholder to the shareholder to compensate for the risk within the insurance contract that has been conferred upon the supporting insurance risk capital.
- *Tax compensation*: a cashflow from policyholder to shareholder to compensate for an additional layer of tax on investment income on insurance risk capital that arises within the insurer.
- *Contractual service margin*: a cashflow from the policyholder to service providers to reward utility received from the services provided under the insurance contract and includes a return on operational risk capital allocated to the contract.

Along with other fulfilment cashflows, these three elements of profit also need to be measured for appropriate recognition against insurance risk capital and services.

Under IFRS17, the total premium (cashflow B in the diagram above) and the present value of all cashflows specific to the insurance contract (cashflows C, D, E, F & G) are essentially deemed to be equivalent at inception. The CSM is the balancing item that makes the cashflows sum to the total premium, so long as it is positive. If it would otherwise be negative, then the contract is considered onerous and a loss is recognised.

2. SUGGESTED APPROACH TO MEASUREMENT OF INSURANCE CONTRACTS

An insurance contract is considered in this paper to have two components that may be measured separately:

- *Insurance risk*: being the exchange of loss cashflows for a premium between policyholder and insurer; and
- *Services*: being services that are delivered to the policyholder for a price. These services could include such categories as distribution, product & underwriting, claim handling, supply chain and corporate.

These two measurement components collectively fulfil the requirements of IFRS17 paragraph 32.

A proposed approach to measurement of an insurance contract is demonstrated in the main body of this paper with a practical example. Cashflows are generated that are then shown in an IFRS17 GMM format. There remain some aspects of the profit and loss, balance sheet and accounting disclosures that are open to interpretation, hence the final approach may differ from what is proposed here.

Insurance contracts may also be measured using the PAA if their coverage period is one year or less. A simplification of the GMM formula proposed above is also provided that complies with the requirements of the PAA. The proposed simplification may be demonstrated to be materially similar to that of the GMM.

Included in Appendix A is the derivation and justification of all formulae used in the practical example.

3. FULFILMENT OF CASHFLOWS FOR THE INSURANCE RISK COMPONENT

3.1. Proposed General Measurement Model for the insurance risk component

The proposed GMM approach that complies with IFRS17 Paragraph 32(a) in respect of the cashflows from the insurance risk component of a contract, from the time of inception until extinguishment, is equivalent to:

- **Present value of uncertain future loss cashflows, discounted at a risk adjusted rate; plus**
- **Present value of the tax on investment income on insurance risk capital, grossed-up for tax and discounted at a 'risk free' rate.**

The 'risk free' rate for the purpose of this paper is defined as that consistent with IFRS17 paragraph 36.

This is the approach for the determination of a fair premium for an insurance risk component of a contract described by (Myers & Cohn, 1981). When certain conditions are met, the Myers & Cohn (**MC**) approach simplifies to the above. This is demonstrated in the Appendix.

3.2. Risk adjusted discount rate

The Capital Assets Pricing Model (**CAPM**) provides a useful and widely understood approach for deriving an appropriate risk adjusted rate to discount cashflows as follows:

$$r_L = r_f + \beta_L(r_m - r_f)$$

Where:

- r_L is the risk adjusted discount rate to apply to uncertain loss cashflows;
- r_f is the risk-free rate, which in the present context allows for the term structure and illiquid nature of the cashflows (IFRS17 paragraph 36);
- r_m is the expected market return; and
- β_L is the CAPM 'Beta' for the uncertain losses.

3.3. Compensation for tax on insurance risk capital

To be at least indifferent about injecting capital into an insurer, shareholders need to be compensated for tax on investment income on insurance risk capital that an insurer will incur. A compensation for tax is therefore included as the second part to the proposed measurement approach.

3.4. Insurance contract assumptions

Assumptions for the insurance risk component of an insurance contract are given in the following table and discussed further below. These assumptions are made to illustrate the measurement of insurance liabilities approach and their derivation is beyond the scope of this paper.

Table A: Insurance assumptions

			Time	*Losses*
Risk free rate	3.0%	p.a.		
Market risk premium	6.0%	p.a.	1	150.00
Loss cashlow β	(0.20)		2	300.00
Tax rate	30.0%		3	200.00
Capital (pre tax effect)	50.0%	of discounted loss reserves	4	100.00
Capital investment β	0.50		5	50.00
Coverage period	1 year			
Derived values				
Risk adjusted rate	1.8%	p.a.		
Capital return rate	6.0%	p.a.		
Capital (post tax effect)	35.0%	of discounted loss reserves		

- A risk-free rate of 3.0% p.a. has been assumed. This is the assumed earning rate for the invested loss reserves and other balances and the risk-free rate for the CAPM calculations of expected returns. The risk-free rate is that inherent in an asset portfolio that replicates the expected cashflows with reference to the term structure and its illiquid nature, but with no risk. This risk-free rate itself requires comprehensive consideration under IFRS17 paragraph 36 which is beyond the scope of this paper. It may be thought of, for example, as a curve of forward discount rates based on sovereign cash and bonds plus an illiquidity premium. However, for the purpose of this paper, this rate will be referred to simply as the 'risk free' rate.
- A market risk premium of 6.0% p.a. has been assumed, which is equal to $(r_m - r_f)$. This will be utilised to assess required discount rates and earning rates under the CAPM framework.
- Loss cashflows are assumed to have a CAPM Beta of -0.2.
- Capital (pre-tax effect) in respect of insurance risk at any point in time is assumed to be 50% of loss reserves in a tax-free environment.
- A single corporate tax rate of 30% is assumed.
- Capital is assumed to be invested in a portfolio that has a CAPM Beta of 0.5.
- A coverage period of 1 year has been assumed for the insurance contract, which is common for a property & casualty contract. The PAA will hence also be illustrated.

- Risk adjusted rate to apply when discounting losses is 3.0% -0.2×6.0% = 1.8%. This hence includes a risk adjustment of -1.2% against the risk-free rate.
- Capital return rate is the annual rate of investment income earned by capital. This is also assessed using the CAPM formula and is equal to 3.0% +0.5×6.0% = 6.0%
- Capital (post tax effect) is the insurance risk capital requirement in a tax-free environment reduced for the corporate tax rate and is hence equal to 35% = 50% × (1 - 30%).
- A total of $800 of losses are expected to arise under the insurance contract, payable at the end of the year in the pattern shown.

3.5. Insurance profit before tax

Using the assumptions above, the measurement of the insurance contract before tax is illustrated in the following table. An explanation for each column is provided. This includes only the loss cashflows and investment income on loss reserves with tax and earnings on insurance risk capital considered in a later table. All cashflows occur at the time indicated which is measured in years from inception of the insurance contract.

Table B: Insurance profit before tax

Time	*Insurance premium (losses only)* (1)	*Loss payments* (2)	*Discounted loss reserves* (3)	*Movement in discounted loss reserves* (4)	*Underwriting profit/(loss)* (5)	*Investment income on loss reserves* (6)	*Insurance profit* (7)	*'Risk free' discount factors* (8)	*Risk adjusted discount factors* (9)
0	765.26		(765.26)	(765.26)	0.00		0.00		
1		(150.00)	(629.03)	136.23	(13.77)	22.96	9.18	0.9709	0.9823
2		(300.00)	(340.35)	288.68	(11.32)	18.87	7.55	0.9426	0.9649
3		(200.00)	(146.48)	193.87	(6.13)	10.21	4.08	0.9151	0.9479
4		(100.00)	(49.12)	97.36	(2.64)	4.39	1.76	0.8885	0.9311
5		(50.00)		49.12	(0.88)	1.47	0.59	0.8626	0.9147

1. The insurance premium for the losses component only is $765.26 and is equal to the discounted value of loss payments at the risk adjusted rate of 1.8%. It is assumed that the premium charged for the insurance risk component of the contract is exactly this figure. Column (9) contains the set of risk adjusted discount factors such that the premium is equal to columns (2)×(9).
2. Loss payments that are made at the end of each year per the assumptions.
3. Discounted loss reserves are equal to the present value of future loss payments discounted at the risk adjusted rate. This is also equal to columns (2)×(9), but just for the loss payments expected in future years.
4. Movement in discounted loss reserves is the annual movement of column (3).
5. Underwriting profit/(loss) is equal to (1)+(2)+(4). This is a traditional measure of insurance contract performance.

6. Investment income on loss reserves is equal to 3% multiplied by the opening balance for the year of discounted loss reserves from column (3).
7. Insurance profit is equal to (5) + (6).
8. Risk free discount factors, calculated here for use throughout the worked example and equal to $(1+3.0\%)^{-Time}$.
9. Risk adjusted discount factors, calculated here for use throughout the worked example and equal to $(1+1.8\%)^{-Time}$.

The insurance profit that emerges each year is equivalent to the risk adjustment, being the negative of the Beta of liabilities times the market risk premium, (-1)×(-0.2)×6% = 1.2%, multiplied by the opening loss reserves.

3.6. Capital, Tax and Tax investment income compensation

Thus far, the operation of the risk adjustment has been demonstrated for losses. In this section, tax is considered and the compensation for tax that is incurred on expected investment returns on insurance risk capital.

Table C: Capital, Tax and Tax on investment income compensation

Time	*Insurance risk capital (reduced for tax effect)* (10)	*Investment income on capital* (11)	*Tax balance capital investment income* (12)	*Movement in tax balance* (13)	*Investment income on tax balance* (14)	*Total profit before tax* (15)	*Tax* (16)	*Total profit after tax* (17)	*Return on capital* (18)
0	267.84		(16.38)	(16.38)		0.00	0.00	0.00	
1	220.16	16.07	(9.98)	6.40	0.49	32.14	(9.64)	22.50	8.40%
2	119.12	13.21	(4.62)	5.36	0.30	26.42	(7.93)	18.49	8.40%
3	51.27	7.15	(1.70)	2.92	0.14	14.29	(4.29)	10.01	8.40%
4	17.19	3.08	(0.43)	1.27	0.05	6.15	(1.85)	4.31	8.40%
5		1.03		0.43	0.01	2.06	(0.62)	1.44	8.40%

10. Insurance risk capital (reduced for tax effect) is 35% multiplied by the loss reserves discounted at the risk adjusted rate from column (3).
11. Investment income on capital is equal to the capital at the beginning of the year in column (10) multiplied by the expected earning rate for capital investments of 6%.
12. Tax balance for capital investment income is the compensation that is charged to the policyholder as an additional premium component and released to compensate for tax on the investment income. This is equal to the present value using the risk-free discount rates of the expected future tax payable on the investment income on insurance risk capital, grossed up for tax. Hence it is equal to columns (11) × 30%÷(1-30%) × (8).
13. Movement in tax balance equals the annual movement of column (12).

14. Investment income on tax balance equals the risk-free rate of 3% multiplied by the opening tax balance for the year from column (12).
15. Total profit before tax equals the insurance profit plus investment income on insurance risk capital plus the movement in the tax balance plus investment income on the tax balance; hence columns (7)+(11)+(13)+(14).
16. Tax is 30% of column (15).
17. Total profit after tax equals columns (15) + (16).
18. Return on capital equals column (17) divided by the opening balance of capital from column (10).

The policyholder pays both the premium for losses as well as the tax balance in respect of the insurance risk component of the contract, a total of \$781.63 = \$765.26 + \$16.38 (rounding).

The tax compensation, including its investment income and tax of itself, exactly equals the tax on insurance risk capital investment income each year. That is 30% × (11) = (1-30%) × [(13) + (14)].

The return on insurance risk capital after tax each year is equal to 6.0% from capital investment income, with the tax on this income compensated for, plus the risk adjustment. Capital being 50% of loss reserves before the tax effect means that the annual 1.2% risk adjustment in loss reserves becomes 1.2% ÷ 50% = 2.4%. Hence the total return on capital is 8.4% and this emerges each year provided that, as per this worked example, capital is a constant proportion of loss reserves discounted at the risk adjusted rate for the duration of loss cashflows.

3.7. Internal Rate of Return

The reconciliation of the internal rate of return (**IRR**) for the insurance risk capital supporting the insurance risk component of the contract is shown in the following table.

Table D: Shareholder IRR

Time	*Movement in capital* (19)	*Total profit after tax* (20)	*Shareholder cashflow* (21)
0	(267.84)	0.00	(267.84)
1	47.68	22.50	70.18
2	101.04	18.49	119.53
3	67.86	10.01	77.86
4	34.08	4.31	38.38
5	17.19	1.44	18.63
Internal rate of return			8.40%

19. Movement in capital is the annual change in the insurance risk capital requirement shown in column (10). It reveals the initial injection of capital to support the contract and its subsequent release.
20. Total profit after tax is a repeat of column (17).

21. Shareholder cashflow equals columns (19) + (20), which shows the injection the shareholder makes at inception and the return of that capital plus after-tax profit in subsequent years.

The internal rate of return of column (21) is 8.4%. This is an unsurprising result given the uniform annual return on capital of 8.4% from column (18).

3.8. Reconciliation to the weighted average cost of capital of the insurer

As a small but important aside, it can be shown that the return on insurance risk capital and IRR for this insurance contract of 8.4% p.a. is equivalent to the weighted average cost of capital (**WACC**) of the insurer.

In order to do this, it is assumed that the insurer is in a steady state, continuously writing the one identical insurance contract each year. The balance sheet for this steady state insurer is shown in the following table.

Table E: Steady state balance sheet

Assets			*Beta*
Capital	(22)	675.58	0.50
Loss reserves	(23)	1,930.23	0.00
Tax balance	(24)	33.11	0.00
Liabilities			
Loss reserves	(25)	(1,930.23)	(0.14)
Tax balance	(26)	(33.11)	0.00
Net assets	**(27)**	**675.58**	**0.90**
Weighted average cost of capital	**(28)**		**8.40%**

22. The insurance risk capital of the insurer is equal to the sum of the capital required to support one contract at inception, one at year 1, one at year 2 and so on. Hence the capital required in a steady state of five active contracts is equal to the sum of column (10). This capital is invested in a portfolio with an assumed CAPM Beta of 0.5.
23. Loss reserves in a steady state equal the sum of column (3). As the assets backing the loss reserves are assumed to be invested in risk free assets, the CAPM Beta is zero.
24. The tax balance in a steady state is equal to the sum of column (12). As the assets backing the tax balance are assumed to be invested in risk free assets, the CAPM Beta is zero.
25. Loss reserves in a steady state equal the sum of column (3). Loss reserves have a CAPM Beta of -0.2. As loss payments are tax deductible, the Beta is multiplied by the tax effect of (1-30%) such that the liability Beta on the balance sheet is -0.14.
26. The tax balance in a steady state is equal to the sum of column (12). The tax balance has a CAPM Beta of zero.
27. The net assets have a weighted average Beta of 0.9. This is found by multiplying the Betas by the balance sheet amount, taking the sum and dividing by the net assets.

28. The WACC is then found by applying the CAPM formula to arrive at 8.4% = 3.0% + 0.9×6.0%.

As is revealed in this table, the WACC of the insurer equates to the IRR of the contract within a CAPM framework. Under this framework, an insurer may potentially work backwards from its observed market Beta to determine an overall loss cashflow CAPM Beta for its insurance contracts for use in the risk adjusted rate formula. However, if the insurer undertakes other activities such as insurance services, then this needs to be considered if decomposing an observed CAPM Beta for the insurer to source components.

It was not necessary to assume a steady state; choosing an individual year of a single contract also produces the same result.

4. FULFILMENT OF CASHFLOWS FOR SERVICES AND THE CONTRACTUAL SERVICE MARGIN

4.1. Proposed General Measurement Model approach for services for property & casualty insurers

The proposed measurement that complies with IFRS17 Paragraph 32 in respect of the services provided within an insurance contract, from the time of inception until extinguishment, is equivalent to:

- **Present value of expense cashflows for future services to be provided under the contract, discounted at a risk adjusted rate; plus**
- **A contractual service margin, which equals the restated CSM as of inception plus risk free investment income multiplied by the proportion that the risk (of losses) in any remaining coverage period bears to the total coverage period risk.**

Expense services cashflows may be risk free in the sense that they are contracted to be provided at a certain time for a certain cost. Other services may be subject to risk. For example, claim handling services are relative to the claim frequency and complexity and duration of claim which is often indicated by the average claim size. It is not unusual for property & casualty insurers to assume that claim handling expenses are proportional to losses in quantum and risk.

Five services are defined for the worked example that include distribution, product & underwriting, claim handling, supply chain and corporate. These services have no particular meaning under IFRS17, but they illustrate a range of timing in incurring these expenses either before contract inception, at inception, during the coverage period or beyond the coverage period. The general model invites detailed cashflow projection of service expenses and a release of profit margins that can be attributable to those services proportional to their delivery.

This does not appear to be the case for property & casualty insurers. Under IFRS17, insurers are essentially deemed to provide just the one service under their contracts – coverage of insured events during the coverage period. Profit, being any amount in excess of what is put aside for contract fulfilment, risk adjustment and tax compensation is considered collectively to be CSM relative to the one service of risk coverage. The CSM is therefore released through the coverage period proportional to the risk of insured events, being completely recognised by the end of the coverage period. Perhaps in due course the standards will reflect a release of CSM proportional to the true nature of the services the insurer provides. This may result in profit relating to acquisition expenses being released at inception, a portion of profit during the coverage period and a final portion of profit released in proportion to claim handling and related services.

The CSM cannot be negative. Should it be, then the contract is onerous, and a loss recorded to the extent that the measurement with a zero CSM exceeds the total premium.

If, prior to the end of the contract period, projected fulfilment cashflows change, then the CSM as of inception is reassessed to again be the difference between total premium and the present value of all fulfilment cashflows. The measurement of CSM continues relative to this reassessed figure.

These service modules are typically all necessary for an insurance contract but are not necessarily all provided by the insurer who provides the insurance risk component. A broker or retailing partner may distribute the product for example. The service modules also require operational risk capital to support them. In this worked example, it is assumed that the insurer provides all five service modules to the insurance contract and accounts for them appropriately under IFRS17. Admittedly, the worked example of a 1-year coverage period does not strongly illustrate the CSM concept that would require wholesome consideration for insurance products such as lenders' mortgage insurance.

4.2. Expense assumptions

The expense assumptions for the five services for the worked example are shown in the following table.

Table F: Services expense and profit assumptions

Services	*Expense* (29)	*Profit margin of total premium* (30)
Distribution	100.00	2.50%
Product & UW	40.00	1.00%
Claim handling	5.00% of losses	1.20%
Supply chain	1.00% of losses	1.00%
Corporate	22.50	0.50%
Operational risk capital	45.0% of expenses	

29. The expenses for distribution, product & underwriting and corporate are assumed to be fixed amounts and the service is provided at inception. The two services of claim handling and supply chain are assumed to be provided as losses are paid, proportional to those losses.
30. This is the assumed profit margin percentage of total premium that is charged for the service.

The operational risk capital assumption of 45% of expenses is also expressed prior to the tax effect. Post tax effect the allocated operational risk capital is 31.5% = 45% × (1 - 30%).

4.3. Total premium

With the expense assumptions now defined, the total premium for the insurance contract is shown in the following table.

Table G: Total premium

			Premium	Proportion
Losses				
	PV of losses (risk free)	(31)	743.42	70.43%
Expenses				
	Distribution	(32)	100.00	
	Product & UW	(33)	40.00	
	Claim handling (risk free)	(34)	37.17	
	Supply chain (risk free)	(35)	7.43	
	Corporate	(36)	22.50	
	Total expenses	(37)	*207.11*	19.62%
Profit items				
	Risk adjustment on losses	(38)	21.84	
	PV of tax on income on capital	(39)	16.38	
	Total insurance risk margin	(40)	*38.22*	3.62%
	Risk adjustment on CH & SC	(41)	1.31	0.12%
	Distribution	(42)	26.39	2.50%
	Product & UW	(43)	10.55	1.00%
	Claim handling	(44)	12.67	1.20%
	Supply chain	(45)	10.55	1.00%
	Corporate	(46)	5.28	0.50%
	Contract service margin	(47)	*65.44*	6.20%
	Total profit	(48)	*104.97*	9.94%
Total premium		(49)	**1,055.49**	

31. The present value of losses at the risk-free rate is equal to the losses from column (3) multiplied by the risk-free discount rates in column (8). This is done to show the extent of risk adjustment from discounting at the risk adjusted rate which is shown as a profit item at row (38).
32. Distribution expenses per assumption above, incurred at inception.
33. Product & underwriting expenses per assumption above, incurred at inception.

34. Claim handling is assumed to be 5% of losses and carries the same risk. Shown here is 5% of the risk-free discounted value of losses with the risk adjustment included in row (39).
35. Supply chain is assumed to be 1% of losses and carries the same risk. Shown here is 1% of the risk-free discounted value of losses with the risk adjustment included in row (39).
36. Corporate expenses per assumption above, incurred at inception.
37. Total expenses equal the sum of rows (32) through (36).
38. Risk adjustment on losses. This is the difference in the present value of losses between using the risk adjusted rate and the risk-free rate.
39. Present value of tax on insurance risk capital investment income as calculated earlier in column (12) at time zero.
40. Total insurance risk margin on insurance risk capital equals rows (38) + (39).
41. Risk adjustment on claim handling and supply chain. This is the difference in the present value of claim handling and supply chain expenses between using the risk adjusted rate and the risk-free rate.
42. Distribution profit margin, percentage of total premium.
43. Product & underwriting profit margin, percentage of total premium.
44. Claim handling profit margin, percentage of total premium.
45. Supply chain profit margin, percentage of total premium.
46. Corporate profit margin, percentage of total premium.
47. Contract service margin equals the sum of rows (42) through (46).
48. Total profit equals rows (40) + (41) + (47).
49. Total premium for the insurance contract, which equals rows (31) + (37) + (48).

4.4. Expense reserves and contract service margin

The recognition of expenses and profit on expenses for the service modules are show in the following table.

Table H: Service expenses

Time	Services premium component	Distribution; Product & UW; Corporate	Claim handling	Claim handling reserve	Supply chain	Supply chain reserve	Contract service margin
	(50)	(51)	(52)	(53)	(54)	(55)	(56)
0	273.86	(162.50)		(38.26)		(7.65)	(65.44)
1			(7.50)	(31.45)	(1.50)	(6.29)	
2			(15.00)	(17.02)	(3.00)	(3.40)	
3			(10.00)	(7.32)	(2.00)	(1.46)	
4			(5.00)	(2.46)	(1.00)	(0.49)	
5			(2.50)		(0.50)		

Time	Investment income on services balances	Services profit	Operational risk capital	Investment income on operational risk capital	Total services profit	Tax on total services profit	Total services profit after tax
	(57)	(58)	(59)	(60)	(61)	(62)	(63)
0		0.00	51.19		0.00	0.00	0.00
1	2.36	66.97	2.84	3.07	70.04	(21.01)	49.03
2	1.13	0.45	5.67	0.17	0.62	(0.19)	0.44
3	0.61	0.25	3.78	0.34	0.59	(0.18)	0.41
4	0.26	0.11	1.89	0.23	0.33	(0.10)	0.23
5	0.09	0.04		0.11	0.15	(0.04)	0.10

50. Services premium component is the total charge included in the total premium for the contract services and equal to rows (37) + (41) + (47). In practice, it would typically be calculated backwards from the total premium at row (49) less the insurance risk component that equals the losses (31), the risk adjustment on losses (38) and the tax compensation (39).
51. Distribution, Product & Underwriting and Corporate expenses are assumed to arise at contract inception and are equal to rows (41) + (42) + (45).
52. Claim handling expenses are equal to 5% of loss payments from column (2).
53. Claim handling reserve is 5% of loss reserves from column (3).
54. Supply chain expenses are equal to 1% of loss payments from column (2).
55. Supply chain reserve is 1% of loss reserves from column (3).

56. Contract service margin at inception equals the service premium component from column (50) less the expenses at inception from column (51) less the present value of future service expenses from columns (53) and (55). This balance plus accrued investment income is released over the coverage period proportional to the risk of insured events. In this worked example, the CSM is assumed to be released uniformly over the coverage period of 1 year such that it is zero at time 1.
57. Investment income on services balances equals the risk free rate of 3.0% multiplied by the opening balances from columns (53) and (55) plus one half of the contract service margin in column (56) year 1 only given the assumption of being uniformly released during that year.
58. Services profit equals the sum of columns (50) through (57) plus the opening balances from columns (53) and (55).
59. Operational risk capital equals the capital to expense ratio of 45% multiplied by the tax effect of (1-30%) multiplied by expenses incurred in the coming year from columns (51), (52) and (54).
60. Investment income on operational risk capital is equal to the investment return on capital of 6.0% multiplied by the opening balance for operational risk capital from column (59).
61. Total services profit equals columns (58) + (60).
62. Tax on total services profit equals 30% of column (61).
63. Total services profit after tax equals columns (61) + (62).

5. IFRS 17 GENERAL MEASUREMENT MODEL PRESENTATION

5.1. Introduction

This section presents the cashflows of the worked example from Tables A through H in an IFRS 17 format for:

- The Insurance Contracts Liability (**ICL**) which is a balance sheet item that includes the reserves and balances associated with the claim fulfilment cashflows of the insurance contract; and
- The Profit & Loss including the items associated with the insurance contract.

Following this section, the PAA will be explored.

Included as Appendix B to this paper are the Profit & Loss statements from the insurance contract cashflows from Tables A through H, but presented under two alternative accounting bases:

- AASB1023 which is applicable in Australia where a key difference to the IFRS17 GMM is the use of a risk margin that provides a chosen probability of ultimate sufficiency of the reserves; and
- USGAAP which is applicable in the US where the loss reserves are undiscounted.

It may be of use to reference a familiar presentation of the Profit & Loss to appreciate the differences with IFRS17. The difference in profit recognition between the bases will then be summarised in the last section of this paper.

5.2. Summary of Insurance Contract Liabilities, General Measurement Model

The following table brings together the various reserves and balances over the cashflow duration of the insurance contract. In total it forms the 'Insurance Contract Liabilities' item on the balance sheet.

Table I: Summary of Insurance Contract Liabilities, General Measurement Model

			Time		
Loss reserves	*0*	*1*	*2*	*3*	*4*
Undiscounted loss reserve	(800.00)	(650.00)	(350.00)	(150.00)	(50.00)
Risk free discount	56.58	34.28	15.81	5.78	1.46
Risk adjustment	(21.84)	(13.31)	(6.16)	(2.26)	(0.57)
Discounted loss reserve	(765.26)	(629.03)	(340.35)	(146.48)	(49.12)
Tax compensation balance					
Undiscounted tax compensation balance	(17.37)	(10.48)	(4.82)	(1.76)	(0.44)
Risk free discount	0.99	0.50	0.20	0.06	0.01
Risk adjustment	0.00	0.00	0.00	0.00	0.00
Discounted tax compensation balance	(16.38)	(9.98)	(4.62)	(1.70)	(0.43)
Claim handling reserves					
Undiscounted claim handling reserve	(40.00)	(32.50)	(17.50)	(7.50)	(2.50)
Risk free discount	2.83	1.71	0.79	0.29	0.07
Risk adjustment	(1.09)	(0.67)	(0.31)	(0.11)	(0.03)
Discounted claim handing reserve	(38.26)	(31.45)	(17.02)	(7.32)	(2.46)
Supply chain reserves					
Undiscounted supply chain reserve	(8.00)	(6.50)	(3.50)	(1.50)	(0.50)
Risk free discount	0.57	0.34	0.16	0.06	0.01
Risk adjustment	(0.22)	(0.13)	(0.06)	(0.02)	(0.01)
Discounted supply chain reserve	(7.65)	(6.29)	(3.40)	(1.46)	(0.49)
Insurance Contract Liabilities					
Undiscounted ICL	(865.37)	(699.48)	(375.82)	(160.76)	(53.44)
Risk free discount	60.97	36.84	16.96	6.19	1.56
Risk adjustment	(23.15)	(14.11)	(6.53)	(2.40)	(0.61)
Discounted ICL	(827.55)	(676.76)	(365.40)	(156.96)	(52.49)
Contract service margin	(65.44)				
Total ICL	**(892.99)**	**(676.76)**	**(365.40)**	**(156.96)**	**(52.49)**
Summary of risk adjustment discount					
Undiscounted risk adjustment	(24.55)	(14.82)	(6.82)	(2.49)	(0.62)
Risk free discount	1.40	0.71	0.29	0.09	0.02
Discounted risk adjustment	(23.15)	(14.11)	(6.53)	(2.40)	(0.61)

There is quite an amount of necessary detail here to capture all items of loss reserve, tax balance, risk adjustments, service expenses and contract service margin. However, many of the blocks of information above are all based on simple multiples of the loss reserve cashflows given the nature of the assumptions made. For example:

- The discounted tax compensation balance is proportional to the loss reserve risk adjustment; and
- All cashflows for expenses and the contract services margins are percentages of the loss reserve cashflows.

Expressing the risk adjustment as the difference between using a risk adjusted discount rate and the risk-free rate reveals a risk adjustment that technically has no cashflow. Shown at the bottom of Table I above is a constructed risk-free discount unwind for the risk adjustment for use in IFRS17 reporting. This is found by working backwards from the projected last risk adjustment calculation and adding in the unwind each year.

5.3. Profit & Loss, General Measurement Model

Utilising the cashflows developed previously, the IFRS17 profit and loss that brings together the insurance risk and services components is assembled below:

Table J: IFRS 17 Profit & Loss, General Measurement Model

	Time					
	0	*1*	*2*	*3*	*4*	*5*
Insurance Revenue	**1,055.49**					
Insurance service expense	**(1,055.49)**	**82.06**	**13.66**	**7.39**	**3.18**	**1.07**
Losses paid		(150.00)	(300.00)	(200.00)	(100.00)	(50.00)
Undiscounted loss reserve mvt	(800.00)	150.00	300.00	200.00	100.00	50.00
Discount mvt on loss reserves	56.58	(22.30)	(18.47)	(10.03)	(4.33)	(1.46)
Discount unwind		22.30	18.47	10.03	4.33	1.46
Risk adjustment mvt	(21.84)	8.53	7.15	3.90	1.69	0.57
Undiscounted tax compensation mvt	(17.37)	6.89	5.66	3.06	1.32	0.44
Discount mvt on tax compensation	0.99	(0.49)	(0.30)	(0.14)	(0.05)	(0.01)
Discount unwind		0.49	0.30	0.14	0.05	0.01
Distribution	(100.00)					
Product & UW	(40.00)					
Corporate	(22.50)					
Contract service margin mvt	(65.44)	65.44				
Claim handling		(7.50)	(15.00)	(10.00)	(5.00)	(2.50)
Undiscounted claim handling mvt	(40.00)	7.50	15.00	10.00	5.00	2.50
Discount mvt on claim handling	2.83	(1.12)	(0.92)	(0.50)	(0.22)	(0.07)
Discount unwind		1.12	0.92	0.50	0.22	0.07
Risk adjustment mvt	(1.09)	0.43	0.36	0.19	0.08	0.03
Supply chain		(1.50)	(3.00)	(2.00)	(1.00)	(0.50)
Undiscounted supply chain mvt	(8.00)	1.50	3.00	2.00	1.00	0.50
Discount mvt on supply chain	0.57	(0.22)	(0.18)	(0.10)	(0.04)	(0.01)
Discount unwind		0.22	0.18	0.10	0.04	0.01
Risk adjustment movement	(0.22)	0.09	0.07	0.04	0.02	0.01
Total risk adjustment unwind		0.69	0.42	0.20	0.07	0.02
Insurance service result	**0.00**	**82.06**	**13.66**	**7.39**	**3.18**	**1.07**
Insurance finance income/ (expense)		**(24.83)**	**(20.30)**	**(10.96)**	**(4.71)**	**(1.57)**
Investment income		**44.95**	**33.68**	**18.45**	**8.01**	**2.72**
Loss reserves + balances		25.81	20.30	10.96	4.71	1.57
Shareholder capital		19.14	13.38	7.49	3.30	1.14
Investment result		**20.12**	**13.38**	**7.49**	**3.30**	**1.14**
Profit before tax	**0.00**	**102.18**	**27.04**	**14.88**	**6.48**	**2.21**
Income tax	**0.00**	**(30.66)**	**(8.11)**	**(4.46)**	**(1.95)**	**(0.66)**
Profit for the period	**0.00**	**71.53**	**18.93**	**10.42**	**4.54**	**1.55**

As required under IFRS17, no profit is recognised at the inception of the insurance contract. The mechanism of the risk adjustment, tax compensation and contractual service margin recognises profit proportional to the provision of insurance risk capital by the shareholder and the provision of services.

One of the key aspects in which the IFRS17 Profit & Loss differs from other standards that report an 'underwriting result' is that the discount unwind is credited to the 'Insurance service result' with an exact offset revealed as an 'Insurance finance expense'. A similar outcome would be achieved within current accounting standards by moving investment income on technical provisions into the underwriting result. This change means that the investment result is therefore the sum of:

- Investment income on capital; plus
- Investment income on the CSM; plus
- Market risk included in the investment strategy for assets supporting the 'Insurance Contract Liabilities' (which is zero under the assumptions of the worked example in this paper).

6. IFRS 17 PREMIUM ALLOCATION APPROACH

6.1. Proposed simplification of measurement for the PAA

The PAA may be applied to insurance contracts that have a coverage period of one year or less. At a reporting date, it enables insurance contracts that are currently within their coverage period to measure liabilities associated with the remaining coverage as a proportion of the premium less acquisition costs. The proportion represents the risk (of losses) in the remaining coverage period relative to the total coverage period risk. The PAA essentially replaces the concept of CSM.

Even if the PAA is adopted, incurred claims are measured in the same manner as they are under the GMM. Despite this, the opportunity is taken here to propose a simpler measurement approach for incurred claims that complements the PAA and reduces the extent of detailed cashflows that the GMM invites.

The proposed simplified measurement of *incurred claims only* that complies with IFRS17 Paragraph 32 for an insurance contract, including both insurance risk and services components, is:

- **Present value of uncertain future loss cashflows, discounted at a 'risk free' rate; plus**
- **Present value of claim handling and other service expenses, discounted at a 'risk free' rate; plus**
- **An adjustment for risk.**

The 'risk free' rate is defined as that consistent with IFRS17 paragraph 36.

6.2. Adjustment for risk

A proposed adjustment for risk combines the concepts of a risk adjustment on all cashflows and tax compensation on investment income on insurance risk capital into the one calculation as follows:

- Present value of future loss cashflows discounted at the risk adjusted rate less the present value of loss cashflows discounted at the risk-free rate; *multiplied by*
- An *RA modifier* that is equal to:

$$1 + \frac{\tau \kappa r_K}{(1-\tau)\left(r_f - r_L\right)} + \gamma$$

Where:

- τ is the tax rate, assumed to be 30% in the present worked example;
- κ is the insurance risk capital post tax effect expressed as a ratio of loss reserves discounted at the risk adjusted rate, assumed to be 35%;
- r_K is the expected return on capital, assumed to be 6.0%;
- r_f is the risk-free rate, assumed to be 3.0% and in the present context allows for the term structure and illiquid nature of the cashflows (IFRS17 paragraph 36);
- r_L is the risk adjusted discount rate to apply to uncertain loss cashflows, assumed to be 1.80%;
- γ is the claim handling and supply chain service expense allowance, expressed as a proportion of loss payments and is assumed to be 6.0%; and

In the present example, the RA modifier is equal to:

$$1 + \frac{0.30 \times 0.35 \times 0.06}{(1-0.30) \times (0.030 - 0.018)} + 0.06 = 1.810$$

It is a condition of the RA modifier that the insurance risk capital allocated to the insurance contract is maintained as a constant proportion of the loss reserves discounted at the risk adjusted rate from inception of the insurance contract until the loss cashflows cease.

6.3. Insurance contract liabilities, Premium Allocation Approach

Using the proposed measurement approach under the PAA, the insurance contract liabilities are given in the following table.

Table K: Summary of Insurance Contract Liabilities, Premium Allocation Approach

	Time				
Loss reserves	*0*	*1*	*2*	*3*	*4*
Undiscounted loss reserve	(800.00)	(650.00)	(350.00)	(150.00)	(50.00)
Risk free discount	56.58	34.28	15.81	5.78	1.46
Discounted loss reserves	(743.42)	(615.72)	(334.19)	(144.22)	(48.54)
Claim handling reserves					
Undiscounted projected claim handling	(48.00)	(39.00)	(21.00)	(9.00)	(3.00)
Risk free discount	3.39	2.06	0.95	0.35	0.09
Discounted claim handing	(44.61)	(36.94)	(20.05)	(8.65)	(2.91)
Risk adjustment					
Undscounted risk adjustment	(41.92)	(25.30)	(11.64)	(4.25)	(1.07)
Risk free discount	2.40	1.21	0.49	0.15	0.03
Discounted risk adjustment	(39.53)	(24.09)	(11.15)	(4.09)	(1.04)
Insurance Contract Liabilities					
Undscounted	(889.92)	(714.30)	(382.64)	(163.25)	(54.07)
Risk free discount	62.37	37.55	17.25	6.28	1.57
Discounted loss reserves	(827.55)	(676.76)	(365.40)	(156.96)	(52.49)

The risk adjustment calculation does not readily produce a cashflow pattern and it is not proportional to another cashflow such as loss payments. Hence the risk-free discount and undiscounted risk adjustment need to be determined working backwards from the last risk adjustment to impute a risk-free discount as with the GMM.

6.4. Profit & Loss, Premium Allocation Approach

The IFRS17 Profit & Loss utilising the PAA is as follows.

Table L: IFRS 17 Profit & Loss, Premium Allocation Approach

	Time					
	0	*1*	*2*	*3*	*4*	*5*
Insurance Revenue	**1,055.49**					
Insurance service expense	**(1,055.49)**	**82.06**	**13.66**	**7.39**	**3.18**	**1.07**
Premium less acquisition costs	(892.99)	892.99				
Discount unwind		24.83				
Losses paid		(150.00)	(300.00)	(200.00)	(100.00)	(50.00)
Undiscounted loss reserve movement		(650.00)	300.00	200.00	100.00	50.00
Discount movement on loss reserves		34.28	(18.47)	(10.03)	(4.33)	(1.46)
Discount unwind			18.47	10.03	4.33	1.46
Distribution	(100.00)					
Product & UW	(40.00)					
Corporate	(22.50)					
Claim handling		(9.00)	(18.00)	(12.00)	(6.00)	(3.00)
Undiscounted claim handling movement		(39.00)	18.00	12.00	6.00	3.00
Discount movement on claim handling		2.06	(1.11)	(0.60)	(0.26)	(0.09)
Discount unwind			1.11	0.60	0.26	0.09
Undiscounted risk adjustment movement		(25.30)	13.66	7.39	3.18	1.07
Discount movement on risk adjustment		1.21	(0.72)	(0.33)	(0.12)	(0.03)
Discount unwind			0.72	0.33	0.12	0.03
Insurance service result	**0.00**	**82.06**	**13.66**	**7.39**	**3.18**	**1.07**
Insurance finance income/ (expense)		**(24.83)**	**(20.30)**	**(10.96)**	**(4.71)**	**(1.57)**
Investment income		**44.95**	**33.68**	**18.45**	**8.01**	**2.72**
Loss reserves + balances		25.81	20.30	10.96	4.71	1.57
Shareholder capital		19.14	13.38	7.49	3.30	1.14
Investment result		**20.12**	**13.38**	**7.49**	**3.30**	**1.14**
Profit before tax	**0.00**	**102.18**	**27.04**	**14.88**	**6.48**	**2.21**
Income tax	**0.00**	**(30.66)**	**(8.11)**	**(4.46)**	**(1.95)**	**(0.66)**
Profit for the period	**0.00**	**71.53**	**18.93**	**10.42**	**4.54**	**1.55**

A somewhat simpler Profit & Loss with reduced cashflows yet remaining identical in profit recognition to the GMM under the conditions and assumptions of the proposed measurement approach.

7. PROFIT RECOGNITION PATTERN

7.1. Profit under different accounting standards

A key difference between accounting standards is the pattern in which profit is recognized.

In the table and chart below, the after-tax profit pattern for the insurance contract from the IFRS17 worked example (Tables A through L) is compared with the profit pattern produced under two alternative accounting standards: AASB1023 and USGAAP. The Profit & Loss statements under accounting standards AASB1023 and USGAAP are presented in Appendix B as Tables B1 and B2 respectively.

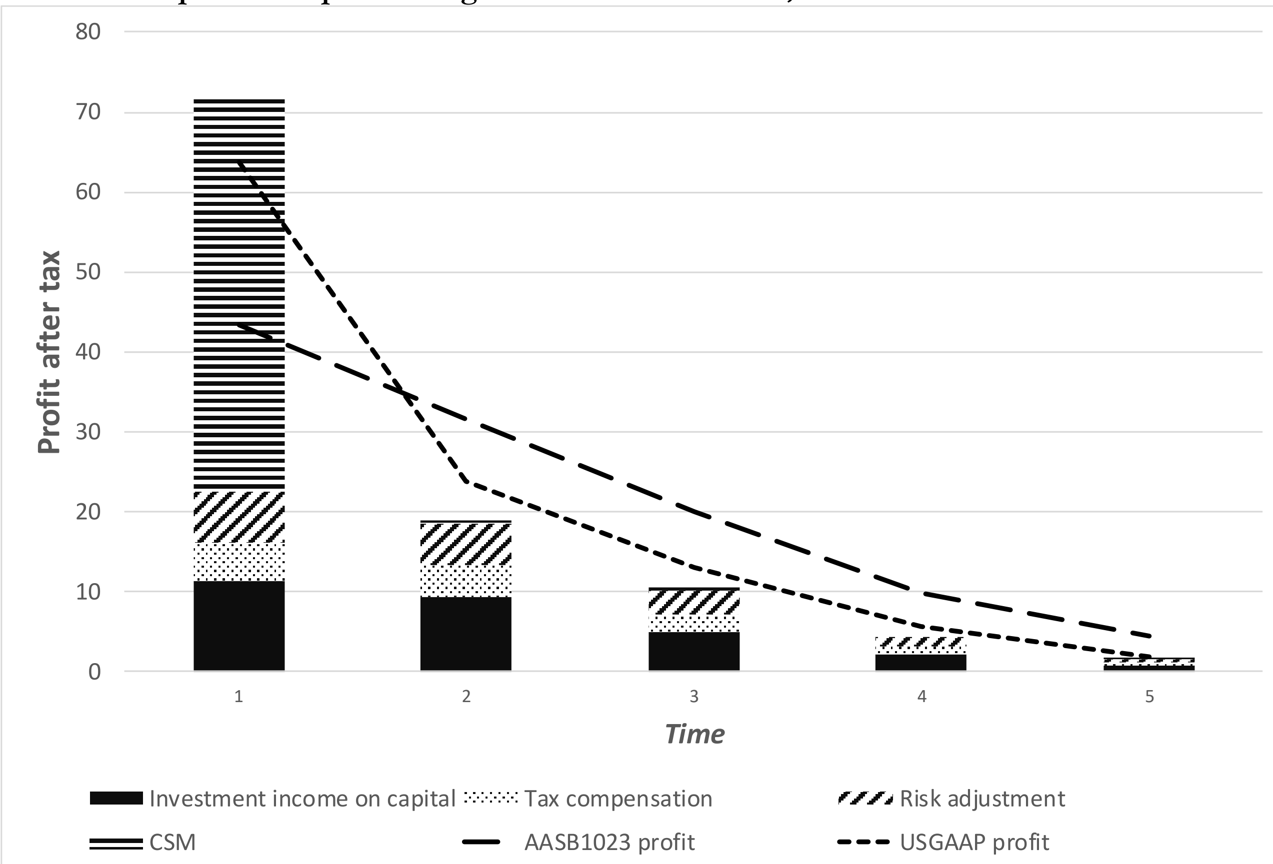
Table M: Recognition of profit after tax

	Profit after tax arising from						
Time	*Investment income on capital* (64)	*Tax compensation* (65)	*Risk adjustment* (66)	*CSM* (67)	*Total profit after tax (IFRS17)* (68)	*Total profit after tax (AASB1023)* (69)	*Total profit after tax (USGAAP)* (70)
1	11.25	4.82	6.43	49.03	71.53	43.39	63.65
2	9.25	3.96	5.28	0.44	18.93	31.63	23.83
3	5.00	2.14	2.86	0.41	10.42	20.08	13.03
4	2.15	0.92	1.23	0.23	4.54	9.73	5.65
5	0.72	0.31	0.41	0.10	1.55	4.51	1.91

64. Investment income on capital after tax, equals column (11) × (1-30%).
65. Tax compensation equals column (11) × 30%.
66. Risk adjustment after tax, equals column (7) × (1-30%).
67. CSM after tax, equal to column (63).
68. Total profit after tax (IFRS17), equals columns (64) + (65) + (66) + (67).
69. Total profit after tax (AASB1023), taken from Appendix B, Table B1.
70. Total profit after tax (USGAAP), taken from Appendix B, Table B2. Note that a 30% tax rate has been used for comparative purposes.

The same information is presented in the following chart.

Chart 1: Comparison of profit recognition between IFRS17, AASB1023 & USGAAP



7.2. Comparison to AASB1023

Industry practice under AASB1023 is to adopt a risk margin as part of total outstanding claim liabilities that provides a chosen probability of sufficiency (**PoS**). The PoS is the proportion of all possible scenario outcomes for future loss payments and associated expenses for which the total outstanding claim liabilities ultimately proves adequate to have met.

The PoS in practice for large listed property & casualty insurers in Australia is typically about 90%. Risk margins of this PoS will generally exceed the cost of capital and service profit margins included in the total premium that delays the recognition of profit relative to the proposed IFRS17 measurement approach.

7.3. Comparison to USGAAP

USGAAP measures loss and loss adjustment expense reserves on an undiscounted basis. This means the profit that is deferred beyond the coverage period is the investment income on the reserves with the underwriting result expected to be zero through the runoff of loss and loss adjustment expense cashflows. Given the prevailing low levels of risk-free investment return, the USGAAP approach currently releases profit somewhat similarly to IFRS17. The risk adjustment and tax compensation under IFRS17 defers profit to a marginally lesser extent than the 3% assumed risk free rate.

8. REFERENCES

Cummins, J. (1990). Multi-period discounted cash flow rate-making models in property-liability insurance. *Journal of Risk and Insurance*, 57(1) 79-109.

Myers, S., & Cohn, R. (1981). *Insurance rate of return regulation and the capital asset pricing model.* Boston, MA: Paper prepared for the Massachusetts automobile rate hearings.

Taylor, G. (1994). Fair premium rating methods and the relations between them. *Journal of Risk and Insurance*, 61(4): 592-615.

Taylor, G. (2004). Economic profit margins in compulsory third party premiums. In *Motor Accidents Authority of NSW: Review of Insurer Profit.* Sydney: The NSW Legislative Council's Standing Committee on Law and Justice 2006.

Turner, A. (1981). Insurance in an Equilibrium Asset-Pricing Model. In J. D. Cummins, & S. Harrington, *Fair Rate of Return in Property-Liability Insurance 1987* (p. Chapter 4). Norwell MA: Kluwer Nijhoff Publishing.

Biography of the Author

Brett Ward is the Chief Actuary for Insurance Australia Group. He is responsible for advising the Board on loss reserves, capital and profit margins. Brett has an economics degree from Macquarie University and is a Fellow of the Institute of Actuaries of Australia.

Appendix A: IFRS17 application of the Myers & Cohn formula

9. CONTEXT

(Myers & Cohn, 1981) develops a formula for determining a fair premium, which is an initial measurement of the cashflows associated with an insurance contract. This formula aligns to the key principles of IFRS17 paragraph 32(a), namely it:

- considers estimates of future cash flows related to the contract;
- adjusts these cashflows to reflect the time value of money and the financial risks related to the future cash flows; and
- includes a risk adjustment for non-financial risk.

This Appendix shows that the proposed approach to the measurement of insurance contracts included in this paper is a special case of the Myers & Cohn (**MC**) formula. There are two conditions needed in order to support a simplified MC formula to measure insurance contracts:

- that there is a single corporate tax rate applied to cashflows of the insurer; and
- the proposed formula and approach to measurement of the insurance contract is applied separately to losses and tax balances until all cashflows associated with the contract cease.

If a third condition is met, it will be also demonstrated that this special case of the MC formula is equivalent to the internal rate of return (**IRR**) approach where the IRR is equivalent to the weighted average cost of capital (**WACC**) within a capital assets pricing model (**CAPM**) framework. This third condition is:

- that insurance risk capital allocated to the contract is a constant proportion of loss reserves discounted at a risk adjusted rate from inception until loss cashflows cease.

In environments where the conditions are not met, then the formulae will not be applicable without adjustment and there may not be equivalence of the IRR to WACC.

10. NOTATION

First, notation is defined for a simplified model of an *insurance risk* contract, meaning the exchange of uncertain losses for a fixed premium.

$t = 0, \ldots, T$	Time periods from the date of contract inception with cashflows generally occurring at integer time t.
P	The insurance risk component premium of the insurance contract.
L_t	Expected losses paid to the policyholder under the insurance contract at time t. Losses are assumed to be payable from time $t = 1, \ldots, T$.
r_f	Risk free rate of return per period.
r_m	Expected rate of return of the market per period.

β_L	The CAPM Beta of losses.
r_L	Risk adjusted discount rate per period applicable to the losses.
K_t	Insurance risk capital requirement, post tax effect, for the insurance contract at time t. Note that K_t may be any reasonable amount relative to the risk of the contract at time t that is consistent with the risk appetite of the insurer. This excludes operational risk capital that supports insurance services.
β_K	The weighted average CAPM Beta of the investments that make up the capital K_t.
r_K	The expected return per period on invested capital.
τ	The corporate tax rate.
λ	The constant proportion of insurance risk capital to loss reserves discounted at the risk adjusted rate – should this approach to capital allocation be adopted.
β_i	The CAPM Beta for the insurer's insurance risk capital, equivalent to the Beta for a share in the insurer that provides contracts with insurance risk only.

Second, notation is added in respect of the *Services* component of the insurance contract.

ρ	Coverage period for the insurance contract that runs from $t = 0 \rightarrow \rho$, where $1 \leq \rho \leq T$.
$\varphi(t)$	Probability distribution function reflecting insured risk within the coverage period, where $\int_{t=0}^{\rho} \varphi(t).dt = 1$.
α	Acquisition service expense amount incurred at inception of the insurance contract.
ξ	Profit margin on acquisition services as a proportion of acquisition service expenses.
γ	Claim handling and other services related to claim fulfilment as a proportion of loss payments.
η	Claim handling and other services profit margin as a proportion of the claim handling and other service expenses.
S_t	The contract service margin at time t.

The generic present value function for cashflows X_t at discount rate r_x is defined as follows:

$$\sum_{y=t+1}^{T} \frac{X_y}{(1+r_x)^{y-t}} = v_t(X; r_x) = v_t^X \tag{1}$$

With the discount rate made clear if the final abbreviated form above is used.

11. FULFILMENT OF CASHFLOWS FOR THE INSURANCE RISK COMPONENT

11.1. Myers & Cohn formula for a fair premium; the initial measurement of an insurance contract

(Myers & Cohn, 1981) proposes the following formula to determine a fair premium that adds two components to the present value of losses to allow for compensation of tax.

$$P = PV(L) + PV(UWPT) + PV(IBT) \tag{2}$$

That is, a fair premium and hence initial measurement of an insurance risk contract at inception equals the sum of:

- Present value of losses (L);
- Tax on the present value of underwriting profit components ($UWPT$); and
- Present value of tax on the investment income of the investible balance of reserves, other balances plus capital (IBT).

The difference in definition between $UWPT$ being the tax applied to present values and the IBT being a present value of tax is significant and implications are explored later.

The notation for the MC formula is abbreviated as follows such that the three premium components can be used more conveniently:

$$P = P^L + P^U + P^B \tag{3}$$

The discount rates used to determine the present value are related to the risk associated with each cashflow. Loss related cashflows are discounted at the risk adjusted rate r_L whereas premium and tax cashflows are discounted at the risk-free rate r_f.

11.2. Present value of losses with risk adjustment for non-financial (insurance) risk

In the absence of tax, the MC approach defines a fair premium as the present value of losses using a risk adjusted discount rate. Using the notation and function defined above, the formula is as follows.

$$PV(L) = P^L = \sum_{t=1}^{T} \frac{L_t}{(1 + r_L)^t} = v_0(L; r_L) \tag{4}$$

The CAPM framework has applicability in providing a suitable discount rate r_L for IFRS17 compliance purposes as follows:

$$r_L = r_f + \beta_L(r_m - r_f) \tag{5}$$

If β_L is negative, meaning that losses have an inverse covariance to market returns, then the discount rate is less than the risk-free rate thus creating a profit for shareholders for the risk they are assuming.

11.3. Compensation adjustments for tax on underwriting and investment income

The underwriting and investment income tax associated with the premium and losses will be dealt with first before considering the tax on investment income on insurance risk capital.

An insurer is (presently) taxed on its underwriting profit, that is, premium less losses for an insurance contract (P-L). Losses are paid over multiple periods and insurers are required to hold reserves for future loss payments. Consequently, premium is recognised as revenue not necessarily at the time of underwriting the contract but in a pattern that is determined by the movement in loss reserves. If we define P^* as the pattern of recognition of premium over multiple periods, then the MC adjustment for underwriting tax may be expressed as follows:

$$PV(UWPT) = P^U = \tau\left[\nu_0\left(P^*; r_f\right) - \nu_0(L; r_L)\right] \tag{6}$$

Note that the star superscript (*) will be used to indicate the recognition pattern for any premium item that appears in the tax adjustments.

An insurer is also taxed on its investment income. Leaving aside insurance risk capital for the moment, investment income is earned on loss reserves, which, for period t, is equal to $r_f \nu_{t-1}(L; r_L)$. If the sum of the underwriting and investment income tax balances is nonzero then this will itself earn investment income and attract tax. It is assumed that the loss reserves are invested in risk free assets replicating the expected loss cashflows, hence the present value of the tax on investment income on loss reserves is:

$$\begin{aligned} PV(IBT) = P^B &= \tau I_0^L + \tau I_0^{UB} \\ &= \tau\left[\sum_{t=1}^{T} \frac{r_f \nu_{t-1}(L; r_L)}{\left(1 + r_f\right)^t}\right] + \tau\left[\sum_{t=1}^{T} \frac{r_f \nu_{t-1}\left(P^{U*} + P^{B*}; r_f\right)}{\left(1 + r_f\right)^t}\right] \end{aligned} \tag{7}$$

Where the notation $\tau I_0^{\bullet}$ refers to the present value at contract inception, discounted at the risk-free rate, of the future expected tax on risk free investment returns for the reserves or balances indicated.

The partially complete MC formula, still excluding consideration of tax of investment income on insurance risk capital, is as follows:

$$P = \nu_0(L; r_L) + \tau\left[\nu_0\left(P^*; r_f\right) - \nu_0(L; r_L)\right] + \tau I_0^L + \tau I_0^{UB} \tag{8}$$

The MC formula can become somewhat iterative, particularly if the underwriting tax rate differs from the tax rate on investment income. The underwriting tax compensation has the property of being equal to the tax of the present value of all amounts recognised in the underwriting result *including* itself. The underwriting and investment income tax balances also generate further investment income and tax. The original MC formula is quite simply stated at equation (2) and, in that form, lends itself to mathematical solving similar to the IRR method. Nevertheless, under the present conditions including a single corporate rate of tax, the formula simplifies.

Separating P^* into its component parts, the MC formula becomes:

$$P = \nu_0(L; r_L) + \tau\left[\nu_0\left(P^{L*}; r_f\right) - \nu_0(L; r_L) + \nu_0\left(P^{U*}; r_f\right) + \tau I_0^{L*} + \tau I_0^{U*}\right] + \tau I_0^L + \tau I_0^{BU} \tag{9}$$

Extract the following three terms from equation (9):

$$\tau\nu_0\left(P^{L*}; r_f\right) + \tau I_0^L - \tau\nu_0(L; r_L) \tag{10}$$

For simplicity, $\nu_t\left(L; r_f\right)$ will also be expressed as ν_t^L, noting the discount rate is r_L. Expanding the terms, equation (10) becomes:

$$\tau\sum_{t=1}^{T}\frac{(\nu_{t-1}^L - \nu_t^L)}{\left(1+r_f\right)^t} \;+\; \tau\sum_{t=1}^{T}\frac{r_f\nu_{t-1}^L}{\left(1+r_f\right)^t} \;-\; \tau\nu_0^L \tag{11}$$

In combining the summations, successive terms for $\nu_1^L \ldots \nu_{T-1}^L$ eliminate leaving the ν_0 terms to also eliminate as follows:

$$\tau\sum_{t=1}^{T}\frac{\left(\nu_{t-1}^L + r_f\nu_{t-1}^L - \nu_t^L\right)}{\left(1+r_f\right)^t} \;-\; \tau\nu_0^L \tag{12}$$

$$\tau\sum_{t=1}^{T}\left[\frac{\nu_{t-1}^L}{\left(1+r_f\right)^{t-1}} - \frac{\nu_t^L}{\left(1+r_f\right)^t}\right] \;-\; \tau\nu_0^L \tag{13}$$

$$\tau\nu_0^L \;-\; \tau\nu_0^L \;=\; 0 \tag{14}$$

Hence:

$$\nu_0\left(P^{L*}; r_f\right) - \nu_0(L; r_L) = -I_0^L \tag{15}$$

This indicates that the investment income at the risk-free rate appears as a negative item in the underwriting tax compensation. This is a tax deduction in the form of the risk-free discount 'unwind' on the loss reserves that appears in the underwriting result.

Extracting the underwriting tax term from equation (9):

$$P^U = \tau\left[\nu_0\left(P^{L*}; r_f\right) - \nu_0(L; r_L) + \nu_0\left(P^{U*}; r_f\right) + \tau I_0^{L*} + \tau I_0^{BU*}\right] \tag{16}$$

Substituting in equation (15):

$$P^U = \tau\left[-I_0^L + \nu_0\left(P^{U*}; r_f\right) + \tau I_0^{L*} + \tau I_0^{BU*}\right] \tag{17}$$

Equation (17) is solved with $P^U = -\tau I_0^L$ such that $P^B = \tau I_0^L$ and the tax balances sum to zero, thus:

$$P^U = \tau[-I_0^L - \tau I_0^{L*} + \tau I_0^{L*}] = -\tau I_0^L \tag{18}$$

The MC formula then simplifies to:

$$P = \nu_0(L; r_L) - \tau I_0^L + \tau I_0^L \tag{19}$$

$$P = \nu_0(L; r_L) \tag{20}$$

Simply the present value of losses discounted at the risk adjusted rate.

While it seems considerable effort to demonstrate that the MC formula tax compensations include an underwriting tax discount unwind deduction that offsets the tax payable on the investment income for loss reserves; what is not compensated for is also of importance.

At first glance, it appears as though the MC formula would compensate for <u>all</u> tax payable by the insurer; that the total profit from an insurance contract is the same as if no tax applied. It is common to read comments about the MC approach, including from Myers & Cohn themselves, that the fair premium includes the *present value of the tax burden on the insurer's underwriting and investment income.* However, this is not quite the case. The MC formula does not compensate shareholders in respect of all tax compared to a tax-free environment. There are two aspects of profit for which shareholders are not compensated for:

- *Insurance risk adjustment.* As losses are discounted at the risk adjusted rate in the underwriting tax compensation calculation, the risk adjustment emerges in the underwriting result without compensation and is therefore taxed. This is appropriate. As losses are tax deductible, the taxing authority is in effect a quota share partner and hence receives the appropriate share of the risk adjustment. As the net risk for the shareholder has been reduced, the risk adjustment net of tax is an appropriate return and no compensation from the policyholder is required.
- *Investment risk on loss reserves and tax balances.* Insurers often leverage investment risk on the loss reserves and other balances, particularly credit risk and duration mismatch risk. The risk, profit and tax are all borne by the shareholder and no compensation from the policyholder is required.

11.4. Compensation for tax on investment income on insurance risk capital

The taxing of investment income earned on capital represents an extra layer of taxation. Shareholders could otherwise invest the capital directly into the investments made by the insurer. By investing in an insurer, the shareholder experiences another layer of tax on investment income received by the insurer before it is ultimately distributed to the shareholder. In order to be at least indifferent about investing in an insurer to enable insurance contracts to be underwritten, shareholders need to be compensated for this extra layer of taxation.

Using the present notation and the assumption that cashflows occur at the end of each period t, the insurance risk capital that is required during period t that is required to support the uncertain loss reserves is equal to K_{t-1} which is the capital at the start of and for the duration of period t. This is assumed to have an expected return in each period of $r_K = r_f + \beta_K(r_m - r_f)$.

The tax compensation required in each period t is equal to $\tau r_K K_{t-1}$. The balance that delivers this compensation will also be subject to underwriting tax as well as income tax on investment earnings and needs to be 'grossed up' to allow for this. Although given the single corporate tax rate, this could be done intuitively by taking the present value of $\tau r_K K_{t-1}$ and using a $1/(1-\tau)$ gross up factor. A more detailed approach is taken here, including showing the MC formula components.

Insurers will have a strategic asset allocation to apply to its capital that will have a market Beta, β_K. The MC formula is agnostic to the Beta of capital and compensates the actual extent of the extra layer of taxation despite this being expectedly higher the higher the Beta. There is hence no strategic asset allocation considered other than fair by the MC formula.

Commencing with the last period of cashflows that occur at time T, the tax balance at the start of the period plus investment income needs to exactly deliver:

- underwriting tax on the part of the balance recognised in the period, which is all of the remaining balance for the last period T;
- tax on investment income earned by the balance; and
- the actual compensation on the tax on insurance risk capital investment income.

This is shown in equation (21). To simplify notation, the present value (risk free rate) at time t of the future tax balance cashflows in respect of the tax on the investment income on investment risk capital is denoted as v_t^θ.

$$v_{T-1}^\theta(1 + r_f) - \tau v_{T-1}^\theta - \tau r_f v_{T-1}^\theta - \tau r_K K_{T-1} = 0 \quad (21)$$

Which can be arranged as follows:

$$v_{T-1}^\theta(1 + r_f)(1 - \tau) = \tau r_K K_{T-1} \quad (22)$$

$$\nu_{T-1}^{\theta} = \frac{\tau r_K K_{T-1}}{(1-\tau)\left(1+r_f\right)} \tag{23}$$

The preceding period is then added, which is in the same form as equation (21), except that:

- The underwriting tax is on the movement in the balance that occurs during period $T\text{-}1$; and
- The outcome of the opening tax balance together with the period $T\text{-}1$ cashflows needs to deliver the required tax balance at the end of the period.

Hence the tax balance activity for period $T\text{-}1$ may be expressed as:

$$\nu_{T-2}^{\theta}\left(1+r_f\right) - \tau\left(\nu_{T-2}^{\theta} - \nu_{T-1}^{\theta}\right) - \tau r_f \nu_{T-2}^{\theta} - \tau r_K K_{T-2} = \nu_{T-1}^{\theta} \tag{24}$$

Then, rearranging to determine ν_{T-2}^{θ}:

$$\nu_{T-2}^{\theta}\left(1+r_f\right)(1-\tau) = \tau r_K K_{T-2} + \nu_{T-1}^{\theta}(1-\tau) \tag{25}$$

$$\nu_{T-2}^{\theta}\left(1+r_f\right)(1-\tau) = \tau r_K K_{T-2} + \frac{\tau r_K K_{T-1}}{\left(1+r_f\right)} \tag{26}$$

$$\nu_{T-2}^{\theta} = \frac{\tau r_K K_{T-2}}{(1-\tau)\left(1+r_f\right)} + \frac{\tau r_K K_{T-1}}{(1-\tau)\left(1+r_f\right)^2} \tag{27}$$

$$\nu_{T-2}^{\theta} = \frac{\tau r_K}{1-\tau} \sum_{y=T-1}^{T} \frac{K_{y-1}}{\left(1+r_f\right)^{y-(T-2)}} \tag{28}$$

Adding back all preceding periods of t, the tax balance required to include in the premium to compensate for the tax on investment income on insurance risk capital is:

$$\nu_0^{\theta} = \frac{\tau r_K}{1-\tau} \sum_{t=1}^{T} \frac{K_{t-1}}{\left(1+r_f\right)^t} \tag{29}$$

The tax compensation for investment income on insurance risk capital can also be expressed as the present value of the tax payments for inclusion in the MC formula as follows:

$$\nu_0^{\theta} = \tau \sum_{t=1}^{T} \frac{\left(\nu_{t-1}^{\theta} - \nu_t^{\theta}\right)}{\left(1+r_f\right)^t} + \tau \sum_{t=1}^{T} \frac{r_f \nu_t^{\theta}}{\left(1+r_f\right)^t} + \tau r_K \sum_{t=1}^{T} \frac{K_{t-1}}{\left(1+r_f\right)^t} \tag{30}$$

The first term is the underwriting tax that arises as the balance is recognised in the underwriting result. The second term represents the tax on the investment income on the balance itself. The last term is the compensation for the tax on insurance risk capital investment income. These three terms could be added to equation (19) to retain the original form of the MC formula with underwriting and investment income tax components.

The first two terms of equation (30), observing equations (11)-(14) equate to τv_0^{θ}. Thus equation (30) is shown to equate to the same tax balance as shown in equation (29) as follows:

$$v_0^{\theta} = \tau v_0^{\theta} + \tau r_K \sum_{t=1}^{T} \frac{K_{t-1}}{\left(1+r_f\right)^t} \tag{31}$$

$$v_0^{\theta} = \frac{\tau r_K}{1-\tau} \sum_{t=1}^{T} \frac{K_{t-1}}{\left(1+r_f\right)^t} \tag{32}$$

11.5. Simplified MC formula and measurement of insurance contracts

If we bring together equations (20) and (29) then we have the simplified MC formula for determining a fair premium:

$$P = \sum_{t=1}^{T} \frac{L_t}{(1+r_L)^t} + \frac{\tau r_K}{1-\tau} \sum_{t=1}^{T} \frac{K_{t-1}}{\left(1+r_f\right)^t} \tag{33}$$

The formula may also be applied to measure the insurance contract throughout the duration of the loss cashflows. If we define M_t as the measurement of the insurance contract at time t, the formula generalises to:

$$M_t = \sum_{y=t+1}^{T} \frac{L_y}{(1+r_L)^{y-t}} + \frac{\tau r_K}{1-\tau} \sum_{y=t+1}^{T} \frac{K_{y-1}}{\left(1+r_f\right)^{y-t}} \tag{34}$$

To this point, the simplified forms of the MC formula in equations (33) and (34) has been relaxed as to the basis for determining the insurance risk capital requirement K_t. Note that for the next sections dealing with an equivalence of IRR and WACC, the condition of allocated insurance capital being a constant proportion of loss reserves discounted at the risk adjusted rate is asserted.

11.6. WACC for the insurer within a CAPM framework

The balance sheet for an insurer, similar to most companies, may be simply represented as:

$$\text{Capital} = \text{Assets} - \text{Liabilities} \tag{35}$$

Under a CAPM framework, the market Betas weight across the balance sheet as follows:

$$\beta_{Capital}.\text{Capital} = \beta_{Assets}.\text{Assets} - \beta_{Liab}.\text{Liabilities} \tag{36}$$

Under the simplified model of an insurer presented in this Appendix, some of the assets and liabilities on the balance sheet are risk free with a zero Beta. The non-zero Beta asset and liability items on the balance sheet at time t are:

- *Insurance risk capital asset portfolio*. This has the value of K_t that is invested in an asset portfolio with a Beta of β_K.
- *Loss reserve liability*. The loss reserves have a Beta of β_L although the loss reserves are weighted by $(1-\tau)$ to allow for the impact of the tax deduction for losses.

Other items on the balance sheet are assumed to have a zero Beta and these include assets backing loss reserves and tax balance items.

At this point, the condition that the required amount of insurance risk capital is proportional to discounted loss reserves at the risk adjusted rate is asserted as follows:

$$K_t = \lambda(1-\tau)\, v_t(L; r_L) \tag{37}$$

The insurance risk capital factor λ refers to that applicable in a tax-free environment and hence needs to be adjusted for the impact of the corporate tax rate applicable in the environment under consideration. In utilising some of the formulae, it may be more convenient to utilise $\kappa = \lambda(1-\tau)$ where κ is the insurance risk capital post the tax effect and a more observable figure on the balance sheet of an insurer, subject to allowing for operational risk or other capital elements that may form part of capital for an insurer

The Beta of the insurer, β_i, is determined using equation (36) as follows, again making use of v_t^L as being the present value of future loss cashflows discounted at the risk adjusted rate:

$$\beta_i \lambda(1-\tau) v_t^L = \beta_K \lambda(1-\tau) v_t^L - \beta_L (1-\tau) v_t^L \tag{38}$$

$$\beta_i = \beta_K - {}^{\beta_L}\!/_{\lambda} \tag{39}$$

This looks somewhat intuitive. The overall Beta for an insurer is equivalent to the Beta of the assets the capital is invested in, less the liability Beta grossed up for the ratio of loss reserves to capital. However, the ratio λ is not observed on the actual balance sheet owing to tax effects that need to be grossed-up for.

In the way the model has been defined, particularly insurance risk capital being a constant ratio of loss reserves, the Beta for the insurer remains constant through the duration of cashflows for the insurance contract.

The WACC for the insurer may then be found using the standard CAPM formula with β_i.

11.7. Return on capital and the IRR

We now evaluate the profit that is recognised by the proposed measurement of insurance contracts approach.

The first component is straightforward. In each period t, insurance risk capital earns r_K based on an investment allocation with a market Beta of β_K. Tax on insurance risk capital investment income is compensated for by the tax balance included with the premium and released appropriately through the measurement approach.

In addition, the risk adjustment emerges as profit and is subject to tax without compensation from the policyholder. As established earlier, there is no net tax on risk free investment income on loss reserves. In each period t, the risk adjustment hence emerges as follows:

- A change in the loss reserve balance; less
- Losses paid at the end of the period; plus
- Investment income at the risk-free rate on the opening balance of loss reserves.

The after-tax profit emergence of the risk adjustment in any given period t is therefore:

$$\left[(v_{t-1}^L - v_t^L) - L_t + r_f v_{t-1}^L\right](1-\tau) \tag{40}$$

The closing loss reserve balance can be expressed in terms of the opening balance adjusted for the discount rate and losses paid:

$$v_t^L = v_{t-1}^L(1+r_L) - L_t \tag{41}$$

Substituting equation (41) into (40):

$$\left[(v_{t-1}^L - v_{t-1}^L(1+r_L) + L_t) - L_t + r_f v_{t-1}^L\right](1-\tau) \tag{42}$$

Simplifies to:

$$v_{t-1}^L\left(r_f - r_L\right)(1-\tau) \tag{43}$$

Then the return on capital for period t, which is defined here as r_i, is as follows:

$$r_i = r_k + \frac{\nu_{t-1}^L (r_f - r_L)(1-\tau)}{K_{t-1}} \tag{44}$$

$$r_i = r_k + \frac{\nu_{t-1}^L (r_f - r_L)(1-\tau)}{\lambda(1-\tau)\nu_{t-1}^L} \tag{45}$$

$$r_i = r_k + \frac{r_f - r_L}{\lambda} \tag{46}$$

Then solving for β_i: that is implied in r_i using the CAPM formula:

$$r_f + \beta_i(r_m - r_f) = r_f + \beta_K(r_m - r_f) + \frac{r_f - r_f - \beta_L(r_m - r_f)}{\lambda} \tag{47}$$

Hence, we again arrive at:

$$\beta_i = \beta_K - {}^{\beta_L}\!/_{\lambda} \tag{48}$$

Thus, the return on capital equates to WACC and is the same for all periods of t. It follows that the IRR for the insurance contract is also WACC. This is a desirable outcome for a measurement approach and a useful property of allocating insurance risk capital proportional to loss reserves discounted at a risk adjusted rate.

12. SERVICES COMPONENT AND THE CONTRACTUAL SERVICE MARGIN (CSM)

Expenses associated with services provided with an insurance contract and their profit are now added to the measurement formula. Per the notation defined above, expenses for these services have been simply defined as having an acquisition component that arises at the inception of the insurance contract and a component that is proportional to loss payments related to the claim fulfilment services. Each service also has an explicit profit margin that is charged to the policyholder as part of the premium.

Building the service expenses into premium formula (33):

$$\begin{aligned} &Total\ Premium \\ &\quad = \sum_{t=1}^{T} \frac{L_t}{(1+r_L)^t}[1+\gamma(1+\eta)] + \alpha(1+\xi) \\ &\quad + \frac{\tau r_K}{1-\tau}\sum_{t=1}^{T}\frac{K_{t-1}}{(1+r_f)^t} \end{aligned} \tag{49}$$

The expenses and profit associated with the claim fulfilment service are also subject to the risk adjustment under these assumptions.

The contract service margin component at inception is therefore equal to:

$$S_0 = \gamma\eta \sum_{t=1}^{T} \frac{L_t}{(1+r_L)^t} + \alpha\xi \tag{50}$$

IFRS17 requires that the contact service margin be recognized proportional to the risk of insured events during the coverage period; the pattern of risk during the coverage period defined here as $\varphi(t)$. Hence measurement during the coverage period will include a portion of the contract service margin given by:

$$S_t = \max\left\{0 \;\middle|\; S_0'\left(1+r_f\right)^t \int_t^\rho \varphi(t).\,dt \;\; ;\text{for } t < \rho\right\} \tag{51}$$

Where S_0' indicates any reassessment to time t of the initial contract service margin if loss or expense cashflows have been updated to reflect current information. The contract service profit parameters would be adjusted to equate total premium to the present value of cashflows, although cannot be less than zero.

If risk is assumed to be uniform through the coverage period, and the number of periods over which the coverage extends is a small number, then a useful simplification may be:

$$S_t = \max\left\{0 \;\middle|\; S_0'\left(1+tr_f\right)\frac{(\rho-t)}{\rho} \;\; ;\text{for } t < \rho\right\} \tag{52}$$

Hence the measurement of the insurance contract at time t after inception is:

$$M_t = \sum_{y=t+1}^{T} \frac{L_y}{(1+r_L)^{y-t}}[1+\gamma] + \frac{\tau r_K}{1-\tau} \sum_{y=t+1}^{T} \frac{K_{y-1}}{\left(1+r_f\right)^{y-t}} + S_t \tag{53}$$

Under IFRS17, several items within the Profit & Loss and disclosures may require calculation of items utilizing both risk-adjusted and risk-free rates to explicitly identify the risk adjustment.

13. A SIMPLIFICATION OF THE TAX COMPENSATION

A simplification of the tax compensation is possible if the condition that the required amount of insurance risk capital is proportional to discounted loss reserves at the risk adjusted rate is asserted.

Starting again with the base MC formula for premium using simplified notation:

$$v_0(L; r_L) + v_0^\theta \tag{54}$$

The risk adjustment may be separated against the risk-free discount rate as follows:

$$\nu_0(L;r_f) + \left(\nu_0(L;r_L) - \nu_0(L;r_f)\right) + \nu_0^{\theta} \tag{55}$$

Expanding the last term and expressing the insurance risk capital relative to the discounted loss reserves, noting that the $(1-\tau)$ term eliminates:

$$\nu_0(L;r_f) + \left(\nu_0(L;r_L) - \nu_0(L;r_f)\right) + \tau\lambda r_K \sum_{t=1}^{T} \frac{\nu_{t-1}(L;r_L)}{\left(1+r_f\right)^t} \tag{56}$$

If we take the last summation from equation (56):

$$\sum_{t=1}^{T} \frac{\nu_{t-1}(L;r_L)}{\left(1+r_f\right)^t} \tag{57}$$

Then this simplifies as follows:

$$\frac{1}{\left(1+r_f\right)-(1+r_L)}\left[\sum_{t=1}^{T} \frac{\left(1+r_f\right)\nu_{t-1}(L;r_L)}{\left(1+r_f\right)^t} - \sum_{t=1}^{T} \frac{(1+r_L)\nu_{t-1}(L;r_L)}{\left(1+r_f\right)^t}\right] \tag{58}$$

Expanding the summation terms:

$$\frac{1}{r_f - r_L}\left[\begin{array}{c} \dfrac{\left(1+r_f\right)\nu_0(L;r_L)}{\left(1+r_f\right)} - \dfrac{(1+r_L)\nu_0(L;r_L)}{\left(1+r_f\right)} \\ + \dfrac{\left(1+r_f\right)\nu_1(L;r_L)}{\left(1+r_f\right)^2} - \dfrac{(1+r_L)\nu_1(L;r_L)}{\left(1+r_f\right)^2} \\ + \dfrac{\left(1+r_f\right)\nu_2(L;r_L)}{\left(1+r_f\right)^3} - \dfrac{(1+r_L)\nu_2(L;r_L)}{\left(1+r_f\right)^3} \\ + \cdots \end{array}\right] \tag{59}$$

Then, nothing that $(1+r_L)\nu_t(L;r_L) = L_{t+1} + \nu_{t+1}(L;r_L)$:

$$\frac{1}{r_f - r_L}\left[\begin{array}{c} \nu_0(L;r_L) - \dfrac{L_1 + \nu_1(L;r_L)}{\left(1+r_f\right)} \\ + \dfrac{\nu_1(L;r_L)}{\left(1+r_f\right)} - \dfrac{L_2 + \nu_2(L;r_L)}{\left(1+r_f\right)^2} \\ + \dfrac{\nu_2(L;r_L)}{\left(1+r_f\right)^2} - \dfrac{L_3 + \nu_3(L;r_L)}{\left(1+r_f\right)^3} \\ + \cdots \end{array}\right] \tag{60}$$

The successive negative and positive terms for $\nu_1(L; r_L) \dots \nu_{T-1}(L; r_L)$ eliminate leaving behind $\nu_0(L; r_L)$ and losses $L_1 \dots L_T$ discounted at the risk free rate:

$$\frac{\left(\nu_0(L; r_L) - \nu_0(L; r_f)\right)}{(r_f - r_L)} \tag{61}$$

Substituting equation (61) into equation (56), the MC formula simplifies to:

$$\nu_0(L; r_f) + \left(\nu_0(L; r_L) - \nu_0(L; r_f)\right)\left[1 + \frac{\tau\lambda r_K}{r_f - r_L}\right] \tag{62}$$

Adding in service expense and contract service margin, the measurement at time t after inception of the insurance contract is:

$$M_t = \nu_t(L; r_f)[1+\gamma] + \left(\nu_t(L; r_L) - \nu_t(L; r_f)\right)\left[1 + \frac{\tau\lambda r_K}{r_f - r_L} + \gamma\right] + S_t \tag{63}$$

Which appears more in the form of a central estimate including expenses discounted at the risk-free rate together with a risk adjustment and the contract service margin. This particular risk adjustment has three components expressed relative to the risk adjustment on discounted losses:

- The risk adjustment on losses itself; plus
- The tax compensation for investment income on insurance risk capital; plus
- The risk adjustment on claim handling and other expenses.

In the worked example included in the main body of the paper, formula (63) above is utilized with the substitution of $\lambda = {}^{\kappa}/_{(1-\tau)}$ introduced earlier.

Appendix B: Alternative accounting standard Profit & Loss statements

14. INTRODUCTION

This Appendix includes Profit & Loss statements that utilise the worked example cashflows developed in Table A through H in this paper and presents the Profit & Loss statements under alternative accounting standards.

14.1. AASB1023 (Australia)

There is one alternative assumption required in order present the AASB1023 Profit & Loss statement. Practice under AASB1023 is to adopt a risk margin as part of total outstanding claim liabilities that provides a chosen probability of sufficiency (**PoS**). The PoS is the proportion of all possible outcomes for future loss payments and associated expenses for which the total outstanding claim liabilities ultimately proves adequate to have met.

It is assumed that the adopted risk margin is 10% of losses and claim handling and supply chain expenses discounted at the risk-free rate. Capital and investment income on capital has been assumed to be the same as per the IFRS17 worked example for simplicity.

Table B1: Profit & Loss, AASB1023 Approach

	Time					
	0	*1*	*2*	*3*	*4*	*5*
Written premium	1,055.49					
Unearned premium mvt	(1,055.49)	1,055.49				
Earned premium	**0.00**	**1,055.49**				
Losses paid		(150.00)	(300.00)	(200.00)	(100.00)	(50.00)
Discounted loss reserve mvt		(615.72)	281.53	189.97	95.67	48.54
Claim handling		(9.00)	(18.00)	(12.00)	(6.00)	(3.00)
Discounted claim handling mvt		(36.94)	16.89	11.40	5.74	2.91
Discounted risk margin mvt		(65.27)	29.84	20.14	10.14	5.15
Claims expense		**(876.93)**	**10.26**	**9.51**	**5.56**	**3.60**
Distribution	(100.00)					
Product & UW	(40.00)					
Corporate	(22.50)					
Deferred acquisition costs	162.50	(162.50)				
Underwriting expense	**0.00**	**(162.50)**				
Underwriting result	**0.00**	**16.06**	**10.26**	**9.51**	**5.56**	**3.60**
Investment income on technical provisions		26.79	21.54	11.69	5.04	1.70
Insurance profit	**0.00**	**42.85**	**31.80**	**21.20**	**10.60**	**5.30**
Investment income on shareholder's funds		19.14	13.38	7.49	3.30	1.14
Profit before tax	**0.00**	**61.99**	**45.18**	**28.69**	**13.90**	**6.44**
Income tax	**0.00**	**(18.60)**	**(13.55)**	**(8.61)**	**(4.17)**	**(1.93)**
Profit after tax	**0.00**	**43.39**	**31.63**	**20.08**	**9.73**	**4.51**

The risk margin is typically selected to provide a PoS in the range of 75-90%. This results in a risk margin that will typically be in excess of the margins under IFRS17 that results in a slower recognition of profit.

14.2. US GAAP

Under US GAAP, loss and loss adjustment expense reserves are inflated but undiscounted. This creates an implicit margin equivalent to the amount of risk-free discount that would be applicable to cashflows. The lower the prevailing level of risk-free rates, the faster is the recognition of profit under US GAAP.

Capital and investment income on capital has been assumed to be the same as per the IFRS17 worked example for simplicity. The corporate tax rate assumption of 30% has also been maintained for comparative purposes.

Table B2: Profit & Loss, USGAAP Approach

	Time					
	0	*1*	*2*	*3*	*4*	*5*
Premiums written	1,055.49					
Unearned premium mvt	(1,055.49)	1,055.49				
Premiums earned	**0.00**	**1,055.49**				
Losses paid		(150.00)	(300.00)	(200.00)	(100.00)	(50.00)
Undiscounted loss reserve mvt		(650.00)	300.00	200.00	100.00	50.00
Loss adjustment expenses		(9.00)	(18.00)	(12.00)	(6.00)	(3.00)
Undiscounted loss adjustment mvt		(39.00)	18.00	12.00	6.00	3.00
Losses and loss adjustment expenses incurred		**(848.00)**	**0.00**	**0.00**	**0.00**	**0.00**
Distribution	(100.00)					
Product & UW	(40.00)					
Corporate	(22.50)					
Deferred acquisition costs	162.50	(162.50)				
Underwriting expenses	**0.00**	**(162.50)**				
Underwriting result	**0.00**	**44.99**	**0.00**	**0.00**	**0.00**	**0.00**
Investment income on reserves		26.79	20.67	11.13	4.77	1.59
Investment income on shareholder's funds		19.14	13.38	7.49	3.30	1.14
Profit before tax		**90.92**	**34.05**	**18.62**	**8.07**	**2.73**
Income tax		**(27.28)**	**(10.21)**	**(5.59)**	**(2.42)**	**(0.82)**
Profit after tax		**63.65**	**23.83**	**13.03**	**5.65**	**1.91**