The Term Structure of Capital Costs

(Authors’ names blinded for peer review)

February 2019

Abstract

We develop a dynamic profit maximization model for a financial institution with liabilities of varying maturity, and use it for determining the term structure of capital costs. As a key contribution, our theoretical, numerical, and empirical results show that liabilities with different terms are assessed differently, depending on the company’s financial situation. In particular, we find that for a financially constrained firm, value-adjustments due to financial frictions for liabilities in the far future are less pronounced than for short-term obligations, resulting in a strongly downward sloping term structure. Our findings provide guidance for performance measurement in financial institutions.

JEL Classification: C61, G22, G32.
Keywords: risk management; financial friction; profit maximization; capital allocation; NAIC data.
1 Introduction

How do financial institutions discount liability cash flows in the near and distant future? It has long been established that capital costs are an important valuation component for financial institutions facing risk, particularly in the insurance sector (Cummins and Phillips, 2005), and market-consistent valuation frameworks such as IFRS 17 and Solvency II include corresponding “risk margins” for non-replicable risks (Albrecher et al., 2018). Such firm-specific risk penalties arise in financial models with financing frictions (Froot and Stein, 1998; Zanjani, 2002; Bauer and Zanjani, 2016). However, thus far little is known about the effect of the markups on liabilities materializing in the near and far future, that is, about the term structure of these capital costs. This paper closes this gap in literature by analyzing how markups affect liabilities with different maturities, both theoretically and empirically.

Relying on an extension of such risk management models with financial frictions, we devise a theoretical model for the firm-specific term structure of capital costs. We develop our theory in the context of a property and casualty (P&C) insurer, which typically carry business lines that vary in the length of time it takes for claims to be reported and to settle—referred to as short and long tailed business lines. Hence, this industry provides an ideal laboratory setting for our ideas, also since unique data are available due to regulatory reporting requirements. We estimate discount curves (net of discounting at risk-free rates) that depend on firm characteristics. Our key theoretical and empirical finding is that firms that face financial constraints include hefty markups for liabilities in the near future, with a term structure that is rapidly declining. In contrast, well capitalized firms include a relatively modest markup that is less steep over time—and can even be negative and increasing for companies with extremely high capital levels. The key intuition is that due to the generally profitable though risky business, capital costs—that are high for meagerly capitalized firms and modest for well capitalized firms—have a mean-reverting character at the firm level.

Within the theoretical model, we integrate a general P&C loss structure given via so-called loss triangles into a dynamic profit maximization model for an insurer that economizes on different capitalization options similar to that from Bauer and Zanjani (2018). The model is set in an economic environment with financing frictions (Duffie, 2010), and includes both internal and external capital that can be raised at different costs (Brunnermeier et al., 2012). We derive our key equation of the marginal cost of risk from the company’s optimality conditions, along with a rule for the economic allocation of capital to the different lines. In line with Bauer and Zanjani (2018), we find that while the marginal cost takes the conventional form of the value of future liabilities plus allocated capital costs, the company evaluates uncertain liabilities under adjusted probabilities that reflect company effective risk aversion (Froot and Stein, 1998). However, we demonstrate that the adjustment differs for payments due in the next year versus payments in future development years, since the associated probability weights depend on the company’s (expected) financial situation. This difference in treatment for liabilities with different durations implies differences in the assessment of shorter versus longer tailed business lines. In particular, both differences in the

---

1 Solvency II is a directive within the European Union that codifies and harmonizes insurance regulation. International Financial Reporting Standard (IFRS) 17 is the new international accounting standard for insurance contracts, and particularly their valuation. Both emphasize market-consistent valuation principles.

2 Indeed, this is one of the primary aspects addressed in the vast actuarial literature on loss reserving in non-life insurance. We refer to the textbooks by Wüthrich and Merz (2008), Taylor (2012), and Radtke et al. (2016) for details.
loss distribution as well as expected settlement times between business lines will interact with the financial situation of the company to determine their valuations.

We explore this relationship numerically in the context of a calibrated version of our model using financial information from the NAIC database.\footnote{The National Association of Insurance Commissioners (NAIC) is the U.S. standard-setting and regulatory support organization created and governed by the chief insurance regulators from the 50 states, the District of Columbia and five U.S. territories.} We solve our model numerically in a setting with two business lines and two development periods for the long-tailed line (2L2DY). More precisely, we consider a business selling a workers’ compensation insurance as the long-tailed line, where we assume that the losses develop according to a Chain-Ladder model with jointly normal innovations (Mack, 1993), and selling commercial automobile insurance as the short-tailed line. We implement the firm’s profit maximization problem by dynamic programming on a discretized state space. In line with Bauer and Zanjani (2018), we find that the value of the P&C insurer is concave with an optimal point that results from balancing profit expectations and capital costs. However, our differentiation between the long- and short-tailed business lines allows for analyzing the impact of firm capitalization on the optimal line mix. We find that exposure in the long-tailed line—where payments occur further in the future—is relatively higher for financially constrained firms, whereas the opposite is true for the short-tailed line. This is due to the former facing high capital costs in the short-term and lower capital costs in the long term, so that ”delaying” indemnity payments by increasing exposure to the long-tailed line is optimal.

To obtain firm-specific markups empirically, we aggregate the marginal cost equation across lines so that we express aggregate company premiums in terms of expected aggregate discounted liability payments and capital costs—where the expected value features risk adjustment terms due to future capital costs. We can then identify a firm-specific term structure of risk adjustments since different companies have a different mix of business lines, so that the risk adjustment terms will have a different impact on the firm’s right-hand side of marginal cost equation. We rely on a simple version of the term structure specification by Nelson and Siegel (1987), where the parameters depend on firm capitalization as measured by the surplus to asset ratio or the leverage ratio. Our estimation delivers current period industry cost-of-capital, and parameters governing the company-specific term structure as a function of firm characteristics.

Our industry cost of capital figures vary between 7.5% and 13%, depending on the considered year. The term structure of firm-specific markups differs markedly in the financial situation of the firm. More precisely, for a firm with an average capitalization level (equity to asset ratio \( \approx 0.5 \)), its valuation of near-future (due in 1-3 years) liabilities is 20-80% higher than their present values. However, for liabilities that are due much later (in 4-10 years), the markups are around 10-20% or even less. This pattern is similar for poorly capitalized firms, although here short-term markups are even higher and the term structure is even more steeply downward sloping. For firms with high capitalization level, in contrast, we observe much lower markup levels overall—and they may even display markdowns with an upward sloping term structure. That is, extremely well capitalized firms may evaluate near-term liabilities below their expected discounted value, though this markdown subsides for liabilities in the far future. The findings are in line with our theory, and they are robust to the inclusion of additional firm characteristics.
The Term Structure of Capital Costs

Related Literature and Organization of the Paper

Our paper relates to several strands of literature. First, we bring together economic approaches for risk pricing in financial institutions with actuarial loss forecasting methods, for the purpose of deriving the term structure of capital costs in P&C insurance. A seminal contribution with regards to the former literature is Froot and Stein (1998), who present economic foundations for risk pricing in a setting with costly capital. We directly build on the dynamic extension of Froot and Stein’s work by Bauer and Zanjani (2018), where different modes of capitalization are included. The actuarial literature on loss forecasting and claims reserving methods is extensive (Taylor and Ashe, 1983; Wüthrich and Merz, 2008; Radtke et al., 2016, e.g.). We rely on the common Chain-Ladder forecasting approach (Mack, 1993), although generalizations are possible.

The empirical section borrows from the literature on yield curve specification and estimation, particularly from Nelson and Siegel (1987). The drivers for the choice between long- and short tailed lines and associated costs relate to the the finance literature on debt maturity (Custódio et al., 2013; Mian and Santos, 2011; Xu, 2016). Finally, we contribute to the literature on cost of capital estimation in the insurance sector (Cox and Griepentrog, 1988; Cummins and Lamm-Tennant, 1994; Lee and Cummis, 1998; Cummins and Phillips, 2005). In particular, while we rely on a completely different approach, it is comforting that resulting figures fall in the same region as previous estimates.

The remainder of the paper is organized as follows: Section 2 presents the term structure equation derived from a general model of multi-period profit maximization with a general loss structure in a P&C company, an implementation of the model and numerical results; Section 3 presents details of the empirical study on estimating the term structure of capital costs; Section 4 concludes.

2 Multi-Period Profit Maximization with Loss History

2.1 Loss Structure for a P&C Company

Setting up a profit maximization framework for a P&C company requires modeling the asset and the liability sides. For simplicity, we assume the company’s assets bear no risk and that all the uncertainty originates from the liability side, modeled via claim payment amounts.

A P&C company writes new insurance contracts in each of its business lines at the beginning of every year (accident year), during which accidents occur and losses are reported. However, some of the losses are not reported until the next year or even years after the origination of the contract. Furthermore, only a portion of the payments is settled in the accident year, whereas the remainder of the (unrealized) payments will take several years to settle. The lags in reporting and paying losses are accounted for by considering so-called loss development years. Such a loss structure is typically represented via so-called loss triangles, with one triangle recording incurred (reported) losses, and another triangle recording paid losses. To illustrate, in Figure 1 we consider a P&C insurance company with $N$ business lines with corresponding (paid) loss random variables $L_{n,i+j-1}^{(n,i+j-1)}$, with line identifier $n = 1, 2, \ldots, N$, accident year (AY) $i = 1, 2, \ldots, t-d_n, \ldots, t, \ldots$, development year (DY) $j = 1, 2, \ldots, d_n$, and $i+j-1$ being the calendar year (period). For each variable, we only need to identify the development and calendar year and thus drop the accident
The term structure of capital costs

year subscripts for simplicity.

<table>
<thead>
<tr>
<th>AY</th>
<th>DY</th>
<th>1</th>
<th>2</th>
<th>\ldots</th>
<th>d_n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>L_{1}^{(n,1)}</td>
<td>L_{2}^{(n,2)}</td>
<td>\ldots</td>
<td>L_{d_n}^{(n,d_n)}</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>L_{1}^{(n,2)}</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
</tr>
<tr>
<td>\ldots</td>
<td></td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
</tr>
<tr>
<td>t - d_n</td>
<td></td>
<td>L_{1}^{(n,t-d_n)}</td>
<td>\ldots</td>
<td>\ldots</td>
<td>L_{d_n}^{(n,t-1)}</td>
</tr>
<tr>
<td>t - d_n + 1</td>
<td></td>
<td>L_{1}^{(n,t-d_n+1)}</td>
<td>\ldots</td>
<td>L_{d_n-1}^{(n,t-1)}</td>
<td>L_{d_n}^{(n,t)}</td>
</tr>
<tr>
<td>\ldots</td>
<td></td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
</tr>
<tr>
<td>t - 1</td>
<td></td>
<td>L_{1}^{(n,t-1)}</td>
<td>L_{2}^{(n,t)}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td></td>
<td>L_{1}^{(n,t)}</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Loss triangle for a P&C insurer in business line \( n \) with \( t \) accident years and \( d_n \) development years

In the paid loss triangle, for example, \( L_{1}^{(n,1)} \) to \( L_{d_n}^{(n,d_n)} \) denote amounts paid (if positive, or amount received if negative) for insurance sold at the beginning of year 1 in line \( n \). Thus, every year, there are payments for losses incurred in the current year, as well as for losses developed from previous years. Specifically, payments in the same calendar year consist of the diagonal entries in the paid loss triangle. For example, payments in calendar year \( t \) correspond to \( (L_{1}^{(n,t)}, L_{2}^{(n,t)}, \ldots, L_{d_n}^{(n,t)}) \), which are double-boxed inside Figure 1. \( L_{1}^{(n,t)} \) represents losses from the contract sold in period \( t \). Other losses \( (L_{2}^{(n,t)}, \ldots, L_{d_n}^{(n,t)}) \) are developed from previous years’ losses, which are in oval boxes and themselves make up a triangle in Figure 1. We denote this “historical” loss triangle at time \( t - 1 \) as \( \Delta_{(n,t-1)} = \{L_{i}^{(n,t)}, t - d_n + 1 \leq i \leq t - 1\} \), which contains (partial) loss information from \( t - d_n + 1 \) to \( t - 1 \) and is the only source of uncertainty. Denote \( L_{(n,t)} = (L_{2}^{(n,t)}, \ldots, L_{d_n}^{(n,t)}) \) as losses paid in year \( t \) that developed from \( \Delta_{(n,t-1)} \). To account for the loss development in each accident year, it is common to assume that the paid losses triangles have a Markov structure:

\[
P(\Delta_{(n,t)} | \Delta_{(n,t-1)}, \Delta_{(n,t-2)}, \ldots, \Delta_{(n,1)}) = P(\Delta_{(n,t)} | \Delta_{(n,t-1)}).
\]

Also, as is common, we assume independence across accident years. A Markov structure and the independence assumptions together fit most of the loss reserving methods in the P&C industry. It is possible to relax the independence assumption and allow cross-sectional correlations between accident years, at the cost of more complex derivations.

Under independence and Markov assumptions, loss random variables in each accident year are related as follows:

\[
P \left( L_{j}^{(n,t)} | h(L_{1}^{(n,t-j+1:t-1)}), \ldots, h(L_{1}^{(n,t-j+1)}) \right) = P \left( L_{j}^{(n,t)} | h(L_{1}^{(n,t-j+1:t-1)}) \right).
\]

The losses in the \( j^{th} \) development year only depend on the information of the same accident year
and on a function $h$ of loss information on the previous development years. For example, in the most popular stochastic loss reserving method, the so-called Chain-Ladder approach (Mack, 1993), $h$ is the cumulative summation operation:

$$P\left(L_{j}^{(n,t)} \mid h(L_{1:1:j-1}^{(n,t-j+1:t-1)}), \ldots, h(L_{1}^{(n,t-j+1:t)})\right) = P\left(L_{j}^{(n,t)} \mid \sum_{k=1}^{j-1} L_{k}^{(n,t-k)}\right).$$

### 2.2 A Multi-Line Multi-Period Profit Maximization Model

To fully describe the dynamic liabilities that the P&C company faces, we assume the following underwriting process: At the beginning of every period $t$, the insurer chooses to underwrite certain amounts in each line of business and charges premium $p^{(n,t)}$ in return. The underwriting decision corresponds to choosing an exposure parameter $q^{(n,t)}$. The losses will be realized over the development years, but the payments are always contingent on the exposure parameter and paid loss random variables. Also note that in each period, the total indemnity payment includes losses incurred and paid in the current year, as well as losses developed from the past years and to be paid in the current calendar year. Thus, for business line $n$ in period $t$, the indemnity payment can be presented via the following function $I^{(n)}(\cdot)$:

$$I^{(n,t)} = I^{(n,t)}\left(q^{(n,t)}, L_{j}^{(n,t)}, \{Q^{(n,t-1)}, L^{(n,t)}\}_{\text{history}}\right),$$

where we assume $I^{(n,t)}(\{q^{(n,t)}, 0\}, \{Q^{(n,t-1)}, 0\}) = 0$. $Q^{(n,t-1)}$ is the vector of exposure parameters associated with triangle $\Delta^{(n,t-1)}$ and losses $L^{(n,t)}$. In what follows, we will assume that indemnity payments are proportional to the exposure parameters:

$$I^{(n,t)} = q^{(n,t)} \times L_{1}^{(n,t)} + Q^{(n,t-1)} \times L^{(n,t)},$$

but generalizations are possible at the expense of a more cumbersome analysis (Frees, 2017; Mildenhall, 2017). We denote the aggregate period indemnity across business lines by $I^{(t)} = \sum_{n=1}^{N} I^{(n,t)}$.4

The company collects the full premium $p^{(n,t)}$ at the beginning of each period $t$ on each line. The aggregate period premium is $p^{(t)} = \sum_{n=1}^{N} p^{(n,t)}$. The company can raise capital $B^{(t)} \geq 0$ (or shed capital $B^{(t)} < 0$). The cost of raising capital $B^{(t)}$ is $c(B^{(t)})$ if $B^{(t)} \geq 0$. There is no cost of shedding capital, i.e. $c(B^{(t)}) = 0$ if $B^{(t)} < 0$. The company carries over capital $a^{(t-1)}(1 - \tau)$ from the last period, with $\tau$ denoting the unit frictional cost of internal capital. Raising external capital is always marginally more expensive than keeping internal capital, so we always have $c'(\cdot) > \tau > 0$.

Thus, the company’s assets at the beginning of period $t$ are

$$a^{(t-1)}(1 - \tau) + B^{(t)} - c(B^{(t)}) + p^{(t)}.$$}

---

4In the rest of the paper, we use $X^{(t)}$ as the sum across the lines $\sum_{n=1}^{N} X^{(n,t)}$. $X^{(\cdot)}$ is used to represent the line-by-line collection (vector) $(X^{(1,t)}, \ldots, X^{(N,t)})$, and its subset $X^{(m:t)} = (X^{(m,t)}, \ldots, X^{(n,t)})$. $X^{(n,t:t+s)}$ represents a collection of random variables over discrete time $(X^{(n,t)}, X^{(n,t+1)}, \ldots, X^{(n,t+s)})$. 

---
During period $t$, the assets are invested at a fixed annual interest rate $r$. At the end of period $t$, the company pays the aggregate indemnity $I^{(t)}$ from its insurance policies sold in the current period and previous periods. The surplus of assets over aggregate indemnity, denoted by $a^{(t)}$, can then be carried over to period $t+1$. Thus, we have the following law of motion for the company’s capital:

$$a^{(t)} = (a^{(t-1)}(1 - \tau) + B^{(t)} - c_1(B^{(t)}) + p^{(t)})(1 + r) - I^{(t)},$$

(1)

assuming $a^{(t)} \geq 0$. If the company defaults, it pays out all remaining assets to policyholders. The company cannot shed more capital than it has available. Hence, for $a^{(t-1)} \geq 0$, we require that:

$$B^{(t)} \geq -a^{(t-1)}(1 - \tau).$$

(2)

The objective function for each period can be derived using the revenue (premium collected), minus the costs (indemnity, frictional costs on carrying capital, and financing costs). For each period, the expected aggregate indemnity takes the following form:

$$e^{(t)} = E \left[ I^{(t)}1_{\{a^{(t)}\geq 0, \ldots, a^{(t)}\geq 0\}} + (a^{(t)} + I^{(t)})1_{\{a^{(t)}\geq 0, \ldots, a^{(t)}< 0\}} \mid \Delta^{(t-1)} \right].$$

Note that here we write the remaining assets in case of default as $a^{(t)} + I^{(t)} < I^{(t)}$.

Hence, the company’s period profit function $f$ is:

$$f(s_t = \{a^{(t-1)}, Q^{(t-1)}, \Delta^{(t-1)}\}, c_t = \{q^{(t)}, p^{(t)}, B^{(t)}\})$$

$$= (1 + r)p^{(t)} - e^{(t)} - (1 + r)(\tau a^{(t-1)} + c(B^{(t)}))$$

$$= E \left[ 1_{\{a^{(t)}\geq 0, \ldots, a^{(t)}\geq 0\}}\{(1 + r)p^{(t)} - I^{(t)} - (1 + r)(\tau a^{(t-1)} + c(B^{(t)}))\} - 1_{\{a^{(t)}\geq 0, \ldots, a^{(t)}< 0\}}(1 + r)(a^{(t-1)} + B^{(t)}) \mid s_t \right].$$

(3)

The state “$s_t$” contains all the variables that determine the state of the company at the beginning of period $t$. The control “$c_t$” contains all the variables that the company chooses in maximizing the objective function, also at the beginning of period $t$. In particular, both $s_t$ and $c_t$ are predictable with the information from the loss triangle $\Delta^{(t-1)}$. The insurance company’s ultimate objective is to maximize future expected discounted cash flows, which corresponds to the following infinite horizon optimization problem:

$$\max_{c_t} \sum_{t=1}^{\infty} E[\beta t f(s_t, c_t)],$$

(4)

where $\beta = (1 + r)^{-1}$ is the discount factor. The objective function can be equivalently represented as present value of future dividends as follows:

$$\max_{c_t} E \left[ \sum_{t \leq t^*} -\beta^{-1} B^{(t)} - a^{(0)} \right],$$

(5)

where $t^*$ is the time such that $a^{(1)} \geq 0, a^{(2)} \geq 0, \ldots, a^{(t^*-1)} \geq 0, a^{(t^*)} < 0$ (see Appendix A.1 for the proof).

We solve the optimization with constraints (1), (2), a premium function for each line $n$, and a regulatory constraint if needed. For the premium function, we follow Bauer and Zanjani (2018)
and assume that the premium charged for one line is the expected present (actuarial) value of future losses multiplied by a markup function. The present value of future losses for each line at the end of period $t$ can be represented as

$$R^{(n,t)} = \sum_{j=1}^{d_n} \beta^{j-1} q^{(n,t)} L_j^{(n,t+j-1)}.$$ 

The markup function is a (decreasing) function of company risk $\phi$ and size $\theta$, defined as

$$\pi^{(n)} = \pi^{(n)}(\phi, \theta),$$

with the assumption on partial derivatives:

$$\pi_1^{(n)} = \frac{\partial \pi^{(n)}(\phi, \theta)}{\partial \phi} < 0, \text{ and } \pi_2^{(n)} = \frac{\partial \pi^{(n)}(\phi, \theta)}{\partial \theta} < 0,$$

so a company with greater risk and larger size charges a smaller markup over actuarial value.

$\phi$ is a risk metric that measures the risk of a company given its total indemnities and assets. For measuring “risk,” we assume that the policyholders are concerned with the company’s period solvency, so that the risk depends on total indemnities paid $I^{(t)}$ and total end-of-periods assets $S^{(t)}$,

$$\phi = \phi(I^{(t)}, S^{(t)}),$$

where $S^{(t)} = (a^{(t-1)}(1 - \tau) + B^{(t)} - c(B^{(t)}) + p^{(t)}) (1 + r)$. Here, similarly to Bauer and Zanjani (2018), in addition to obvious monotonicity assumptions ($\phi(I, x) \leq \phi(I, y), x \geq y$, and $\phi(X, x) \leq \phi(Y, x), X \leq Y$), we assume scale invariance of the risk metric, i.e. $\phi(aI, ax) = \phi(I, x), a > 0$.

The key example that we will rely on in our numerical applications is the conditional default probability:

$$\phi(I^{(t)}, S^{(t)}) = \mathbb{P}(I^{(t)} > S^{(t)} | \Delta^{(t-1)}).$$

We note that this specification assumes consumers are myopic in that they are only concerned with the coming period—and not necessary the performance of their contract. This may be justified with the assumption that consumers rely on company ratings that obviously do not depend on the term of the obligation.

Altogether, we have the following premium function for line $n$:

$$p^{(n,t)} = \mathbb{E} \left[ \beta R^{(n,t)} | \Delta^{(t-1)} \right] \times \pi^{(n)}(\phi, \theta). \quad (6)$$

According to Bertsekas (1995), the optimization problem (4) is an infinite-horizon discrete-time stochastic optimal control problem, resulting in the following Bellman equation:

**Proposition 2.1. (Bellman Equation).** The Bellman equation for problem (4) reads:

$$V(a^{(t-1)}, Q^{(t-1)}, \Delta^{(t-1)}) = \max_{q^{(t)}, p^{(t)}, B^{(t)}} \mathbb{E} \left[ I_{\{I^{(t)} \leq S^{(t)}\}} \left( p^{(t)} \beta I^{(t)} - \tau a^{(t-1)} - c(B^{(t)}) + \beta V(a^{(t)}, Q^{(t)}, \Delta^{(t)}) \right) - I_{\{I^{(t)} > S^{(t)}\}} \left( a^{(t-1)} + B^{(t)} \right) | \Delta^{(t-1)} \right],$$

where $S^{(t)}$ is a risk metric that measures the risk of a company given its total indemnities and assets.
subject to (1), (2), and (6)

Here the default threshold for the company is $S(t)$. Once the aggregate indemnity is greater than $S(t)$, the company defaults. We do not consider the option of raising emergency capital to save the company as in Bauer and Zanjani (2018), since the focus of this paper is on how loss history, i.e. past exposures $Q^c(.; t-1)$ and losses $\Delta(.; t-1)$, affect the optimal exposure, raising, and allocation decisions. However, incorporating emergency capital is theoretically straightforward. In particular, when incorporating emergency raising capital, we note that the model in Bauer and Zanjani (2018) will be a special case of the general model here with one development year in all business lines, thus effectively reducing the value function to one dimension with $a^{(t-1)}$. In our general setting, with the company having $N$ lines and each line $n$ having $d_n$ development years, there are a total of $1 + \frac{1}{2} \sum_{n=1}^{N} (d_n^2 + d_n - 2)$ state variables.

### 2.3 Term Structure of Capital Costs

We rely on the first order conditions of the Bellman equation in Proposition (2.1) to derive the marginal cost of risk (the proof is provided in Appendix A.1).

**Proposition 2.2.** (Term Structure Equation). We have the marginal cost of risks presented in the form of term structure of capital costs:

$$
\mathbb{E} \left[ \frac{d_n}{\partial a} \sum_{i=1}^{N} \beta^i L_j^{(n,t+j-1)} \mid \Delta^c(.;t-1) \right] * \pi^{(n)} \left( 1 + \frac{1}{2} \sum_{i=1}^{N} \frac{\pi^{(i)}}{\pi^{(n)}} \mathbb{E} \left[ R^{(i,t)} \mid \Delta^c(.;t-1) \right] \right)
$$

$$
= \sum_{s=0}^{d_n - 1} \left( 1 - c'(B(t)) \right) \cdot \left( \mathbb{E} \left[ I_{\{I^{(t)} \leq S(t), \ldots, I^{*(t+s)} \leq S^{*(t+s)} \}} \beta^{s+1} L_s^{(n,t+s)} \mid \Delta^c(.;t-1) \right] \right)
+ \beta \frac{\partial \rho(I^{(t)})}{\partial q^{(n,t)}},
$$

where $w_t = \begin{cases} 
1 + V_1(a^{(t)}, Q^{c(.;t)}, \Delta^{c(.;t)}) & I^{(t)} \leq S^{(t)} \\
\sum_{i=1}^{N} \frac{1}{1 - c'(B(t))} \mathbb{E} \left[ R^{(i,t)} \mid \Delta^c(.;t-1) \right] * \frac{\pi^{(i)}}{\pi^{(n)}} \frac{\partial \rho}{\partial S^{(n,t)}} & I^{(t)} > S^{(t)}
\end{cases}$

$\rho$ is the risk measure associated with the risk metric $\phi$, and $V_1 = \partial V / \partial a$.

$w_t$ is a weighting function similar to the function in Bauer and Zanjani (2018), with

$$
\mathbb{E} \left[ (1 - c'(B(t))) w_t \right] = 1.
$$

The weighting function $w_{t+s} = 1 + V_1(a^{*(t+s)}, Q^{*(c; t+s)}, \Delta^{*(c; t+s)})$ on $I^{(t)} \leq S^{(t)}, \ldots, I^{*(t+s)} \leq S^{*(t+s)}$ from time $t$ to $t+d_n-1$, where $a^{*(t+s)}, Q^{*(c; t+s)}$ etc. are the (stochastic) future state variables under the optimal policy.

The left-hand side of the equation represents the marginal premium income, adjusted down by company size. The first line of the right-hand side of the equation consists of evaluation of future losses for accident year $t$, namely, losses $(L_1^{(n,t)}, L_2^{(n,t+1)}, \ldots, L_{d_n}^{(n,t+d_n-1)})$, which are dependent through a stochastic loss reserving model. The last line of the right-hand side of the equation are about the capital allocation $\frac{\partial \rho(I^{*(t)})}{\partial q^{(n,t)}}$ multiplied by associated capital costs.
The interpretation of the term structure equation is that valuation of liability and capital allocation should be done by line and development year in each line, in contrast to the previous literature, where only by-line allocation is considered. Not only does the equation highlight the importance of loss reserving in pricing P&C insurance, but it also redefines the capital allocation and risk pricing goal in a P&C insurer. We are interested in the term structure of capital costs, namely \( V_1(\alpha^{(t+s)}, Q^{(t+s)}, \Delta^{(t+s)}) \) and how it varies with capitalization level. In what follows, we solve a basic version of our theoretical model numerically, we explore insurer’s decision at optimality, and we analyze how \( V_1 \) behaves in a dynamic setting.

### 2.4 Implementation – Two Lines and Two Development Years

In this section, we provide an implementation of our theory in the previous section in the context of a P&C insurer. Specifically, the P&C insurer has two business lines and two development years on the long-tailed line (2L2DY). We then calibrate and solve for the model using numerical methods.

In 2L2DY, Line 1 is the long-tailed line with development year loss, the paid loss triangle is a 2x1 triangle, as illustrated in Figure 2. Line 2 is assumed to be the short-tailed line with no development years beyond the accident year. The time period equals to \( AY + DY - 1 \). Therefore, at the end of current period \( t \), the insurer faces losses \( L_1^{(1,t)} \) and \( L_2^{(1,t)} \) from its long-tailed line 1, and \( L_1^{(2,t)} \) from its short-tailed line 2. The loss random variables above the solid lines in Figure 2 are realized before \( t \). The grayed-out \( L_1^{(1,t+1)} \) is not a part of the loss triangle and not realized until the end of the next period \( t+1 \), but it is relevant to the premium written for the accident year \( t \) and therefore related to the insurer’s problem.

We simplify notations \( L_j^{(n,t-2+j)} \) to \( L_j^{(n)} \) and \( L_j^{(n,t-1+j)} \) to \( L_j^{(n)} \), shown in Figure 3, as the prime “\( r \)” denotes state variables in the next period. We put three assumptions on loss triangles: (i) Chain-
Ladder in loss development; (ii) conditional normality of loss distribution; (iii) linear correlation between lines. These assumptions make the model tractable and more efficient to calculate the moments of loss random variables in the Bellman equation. Details of three distributional assumptions are presented in Appendix A.2 and useful in developing numerical solutions.

All in all, we solve the following simplified Bellman equation:

$$V(a, q^{(1)}, L_1^{(1)}) = \max_{q'(1), q'(2), p^{(1)}, p^{(2)}, B} \beta \mathbb{E} \left[ 1_{\{I \leq S\}} (S - I) + 1_{\{I \leq S\}} V(a', q'^{(1)}, L'_1^{(1)}) \right] - a - B,$$

where:

$$S = (a(1 - \tau) + B - c_1(B) + p^{(1)} + p^{(2)})e^r \quad a' = S - I.$$ 

For the premium functions, akin to Bauer and Zanjani (2018), we assume the following specification:

$$p_n = \mathbb{E} \left[ \beta R^{(n)} \right] \times \exp \left\{ \alpha_n - \delta_n \mathbb{P}(I > S) - \gamma_n \mathbb{E}[R] \right\}, n = 1, 2$$

where $R^{(1)} = q'(1)L_1^{(1)} + \beta q'(2)L_2^{(1)}$, $R^{(2)} = q'(2)L_1^{(2)}$, and $R = R^{(1)} + R^{(2)}$ representing the aggregate risk in the premium $p_1 + p_2$. Note that the aggregate risk $R$ does not equal to the aggregate indemnity $I$, because of the long-tailed line. $I$ reflects losses to be paid out in the current time period, while $R$ entails risks exposed in one accident year across two periods. $\mathbb{E}[R]$, instead of $\mathbb{E}[I]$ used in the models without development year, reflects the aggregate scale of the insurance business. The motivation for this specification is that policyholders assess company quality via ratings that reflect the default probability, and increasing the scale of insurance business decreases profit margins. We then can specify the premium function as the product of expected present value of future exposed losses and a corresponding markup function.

In the numerical implementation of 2L2DY model, we choose the premium parameters, loss triangle parameters and company level parameters. In this paper, we choose the same premium parameters for both lines. Although a generalization to two distinct sets of premium parameters is possible, it complicates the model solution and may yield results that are difficult to interpret. We leave corresponding extensions for future research. We set the capital costs as $\tau = 0.03$, $c_1^{(1)} = 0.075$, $c_1^{(2)} = 1.0E-10$, and the risk-free interest rate is $r = 0.03$, as in the “base case” scenario in Bauer and Zanjani (2018). The complete set of parameters is listed in Table 1. A solution to the 2L2DY model and corresponding numerical techniques are detailed in Appendix A.3.

### 2.5 Results – Two Lines Two Development Years

Since the value function and optimal policies are functions of three state variables, it is impossible to capture the results in a single graph. Therefore, we graph functions of capital $a$ and previous exposure on the long-tailed line $q^{(1)}$ on the $x$ and $y$ axis, given two extreme levels of the previous shock $L_1^{(1)}$ (large and small). Figure 4 shows the value function and optimal policies under an extremely small previous shock of two standard deviations below the mean at $L_1^{(1)} = 5.0E7$. Figure 8 shows the solution under an extremely large previous shock of two standard deviations above the mean at $L_1^{(1)} = 1.5E8$. We choose extreme levels to documents the effects.
The value function, optimal raising of external capital, and exposure to the short-tailed line all match the corresponding characteristics in Bauer and Zanjani (2018). More precisely, the value function is concave with an optimal capitalization level that economizes on costly external financing, internal capital costs, and an optimal company size. The firm raises capital if it is severely underfunded or sheds capital (pays dividends) if it is severely overfunded, but remains inactive for capitalization levels around the optimal point. The optimal exposure to the short-tailed line is concave and increasing in the capital level, up to a saturation point where costs associated with scale do not warrant further expansion.

However, the solution here additionally provides insights on how previous exposure in the long-tailed line affect the value function and optimal policies. The ridge in Figures 4(a) and 8(a) depict the “optimal capital line”, which connects $a$’s that maximize $V$ under $q^{(1)} \in [0, 2.0]$. To the left of the line, firm value decreases with capital, reflecting the cost of raising external financing; to the right of the line, firm value decreases as the capital level increases, reflecting the cost of carrying internal capital. The optimal capitalization point increases in both the previous exposure $q^{(1)}$ and the loss realization $L^{(1)}_1$. The pattern is also shown in two-dimensional Figures 5(a) and 9(a).

The optimal capital raising decision is slightly more subtle as seen from Figures 4(b) and 8(b). While with a higher previous exposure on long-tailed line, the company will keep more capital and will start raising capital earlier, with lower previous exposure the company will raise more aggressively for low capital levels. This reflects the increased value with limited loss legacy.

The optimal long-tailed line exposure is depicted in Figures 4(c) and 8(c). Interestingly, the optimal exposure is strictly increasing with previous exposure when the capital level is low—which may be counterintuitive at first sight. As the insurer starts out with low capital, it would sell more

<table>
<thead>
<tr>
<th>Premium</th>
<th>Long-Tailed Line</th>
<th>Short-Tailed Line</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$\beta$</td>
<td>-10.0</td>
<td>-10.0</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>-5.0E-10</td>
<td>-5.0E-10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Triangle</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$f$</td>
<td>2.0</td>
<td>N/A</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>2.5E7</td>
<td>N/A</td>
</tr>
<tr>
<td>$\mu_1^{(n)}$</td>
<td>1.0E8</td>
<td>1.0E8</td>
</tr>
<tr>
<td>$(\sigma_1^{(n)})^2$</td>
<td>6.25E14</td>
<td>6.25E14</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Company</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau$</td>
<td>0.03</td>
</tr>
<tr>
<td>$c_1^{(1)}$</td>
<td>0.075</td>
</tr>
<tr>
<td>$c_1^{(2)}$</td>
<td>1E-10</td>
</tr>
</tbody>
</table>

Table 1: Model parameters
The term structure of capital costs

Figure 4: Value function, optimal external capital raising and exposure decision under small previous shock
Figure 5: 2-dimensional representations of value function (small previous shock)

Figure 6: 2-dimensional representations of long-tailed line exposure (small previous shock)

Figure 7: 2-dimensional representations of short-tailed line exposure (small previous shock)
Figure 8: Value function, optimal external capital raising and exposure decision under large previous shock
Figure 9: 2-dimensional representations of value function (large previous shock)

Figure 10: 2-dimensional representations of long-tailed line exposure (large previous shock)

Figure 11: 2-dimensional representations of short-tailed line exposure (large previous shock)
insurance on the long-tailed line, less on short-tailed line and raise external capital. The reason
is that the insurer will be paying 100% of the loss incurred in the short-tailed line, but only a
fraction of the loss incurred in the short-tailed line, while it earns full premium on both lines. As
a result, the long-tailed line offers a relatively attractive source of financing for a firm in need of
funds. In other words, the long-tailed line can serve to gain short-term financing, at lower cost than
raising external capital. This effect is more pronounced for a high previous exposure, since more
short-term financing is needed.

When the capital is over optimal capital line, the optimal exposure to the long-tailed line in-
creases first with previous long-tailed line exposure and then decreases. The need for financing
declines and the insurer’s objective is to balance the current long-tailed line exposure and previous
one. Thus, when the previous exposure is too high, the insurer chooses to decrease the current
exposure to long-tailed line business. In particular, Figure 6(b) shows that while optimal exposure
to line 1 is increasing in previous exposure for low $a$, the relationship inverts for large capital level
$a$. Similarly, for high capital levels, optimal exposure to the long-tailed line becomes flat as shown
in Figure 6(a).

The optimal exposure to the short-tailed line complements with the optimal exposure to the
short-tailed line as evident from Figures 4(d) and 8(d). As a company has more exposure on its
long-tailed line in the last period, it will be responsible for a greater amount of indemnity developed
from the last period. As a result, the insurer will keep more capital and reduce exposure on its
short-tailed line business. When the capital is high, the insurer would increase its short-tailed line
exposure as it complements with decreasing long-tailed line exposure, as we observe at $a = 1E09$
in Figure pairs 6(b) & 7(b), and 10(b) & 11(b). In terms of exposure with respect to capital, we
can see in Figures 7(a) and 11(a) that the optimal short-tailed exposure increases with capital but
becomes flat when the capital is over optimal capital line, similar to the findings in Bauer and
Zanjani (2018). Most evidently in Figure 11(a) when $q^{(1)} = 2$ and capital is low, it is optimal
for the insurer to almost completely shut down the short-tailed line. On the other hand, when the
capital is high, it is optimal for the insurer to almost completely shut down the long-tailed line, as
seen in Figure 10(a).

The value function also provides insight to the term structure of capital costs. According
to the term structure equation in Section 2.3, the gradient of value function with respect to capital
$V_1$ forms the term structure of capital costs. Because of the concavity of the value function with
respect to capital as seen in Figures 5(a) and 9(a), when the company is undercapitalized or its
capital is less than the optimal capitalization level, then $V_1 > 0$ and the company effectively apply
markup to losses in premium pricing. On the other hand, when the company is overcapitalized or
its capital is more than the optimal capitalization level, then $V_1 < 0$ and the company marks down
on losses. Specifically in Figure 9(a), we observe that the value function displays greater concavity
when $q^{(1)}$ is high, thus larger $V_1$, than when $q^{(1)}$ is low. As a result, the term structure of capital
costs is higher for firms that are financially constrained, lower for firms that are better capitalized,
and even negative for firms that are extremely well capitalized.

The natural next steps would be finding the term structure numerically by calculating $V_1$ for
every development year. However, for our 2L2DY implementation, we can only obtain numerical
calculation of $V_1$ for two periods. A sample calculation is shown in Table 2. When the company
begins the period with financial constraint, i.e. with low capitalization and high past exposure,
$V_1$ at the end of the period is positive. For a company that starts with high capital and low past
exposure, $V_1$ at the end of the period is positive. Both $V_1$’s turn closer to zero at the end of period
At the beginning of period $t$ & At the end of period $t$ & At the end of period $t + 1$

Financially constrained & 0.0254 & -0.0170
Well capitalized & -0.0303 & -0.0196

Table 2: Numerical calculation of $V_1$

$t + 1$. From the sample numerical results, we can see a downward sloping $V_1$ for a financially constrained company, who puts more adjustment to the expected losses in the current period than the next period. On the other hand, we can see an upward sloping $V_1$ for the well capitalized company. Alas, the numerical calculation provides very limited information on the term structure, which ideally spans for ten years for a P&C company. If we set out to numerically solve for a 10-year term structure, it requires implementation of full $10 \times 10$ loss triangle. As discussed in Section 2.2, we will face a model with more than 100 state variables and such model is impossible to be solved using dynamic programming approach. Instead, in the next section, we propose an estimation of the term structure of capital costs in an empirical setting. Our empirical findings are in line with our theoretical results.

3 Empirical Study

An insurer relies on premium income, and also capital to support its expected future losses. Capitals are costly to hold, but necessary to obtain whenever the insurance businesses are less profitable and/or the insurer does not do well financially. In calculating how much premium to charge for additional exposure in future losses, a company considers valuation other than simply pricing at actuarial value of expected future losses. The reason is that companies are risk averse about loss shocks that may happen in the future. If such shocks occur, they consume capital quickly and will put the companies at great risk of default and/or out of competition. For most companies, the losses to be paid out in the near future (e.g. 1-3 years) are of the most concern and companies choose to price them higher than their actuarial value, while losses to be paid out in the far future (e.g. 4-10 years) are priced closer to their actuarial value. We are interested in how a P&C company weighs on their valuation of expected future losses at different point of time, and how the valuation changes among companies with different capitalization levels.

Our theoretical finding, i.e. term structure equation (7), shows that the valuation of loss reserve in each line of business involves two discount factors. The first one is a yield curve, in our theory being flat at $r$. The second is a term structure of future losses, described in our theory by $(1 + V_1(a(t), Q^{(t)}, \Delta^{(t)}))$ from $t$ to $t + d_n - 1$. Since $V_1$ is $\partial V/\partial a$ and does not vary by business lines, the term structure summarizes the sensitivity of premium (aggregate/by line) to future losses (aggregate/by line) from development year 1 to $d_n$. Although it makes sense to generalize term structures of future losses by line, in this paper we remain focus on finding a term structure at company level for different years. Our goal is to estimate the sensitivity of premium to the expected future losses and capital in a P&C company in an empirical study.

Now, we start from identifying the variables of interest and back out an equation for estimation. From our theoretical results (7), multiplying exposure $q^{(n,t)}$ to both sides of the equation and summing up by line, we obtain the following equation:
We obtain total premium for all lines plus a fraction on the left-hand-side and valuation of loss reserve plus capital costs on the right-hand-side. The left-hand side of Equation 9 can be identified using net premium written for all lines in company \(i\) in a given year, denoted by \(P_i = \sum_{n=1}^{N} P_{n,i}\), where \(P_{n,i}\) is the net premium written for line \(n\). The capital costs component on the right-hand side is identified using capital/surplus of a company multiplied by unit cost of capital or return-on-capital, which is a parameter to be estimated. We denote surplus for company \(i\) as \(SURP_i\) and unit cost as \(c\).

Next, we wish to identify both the expected future losses component and term structure component \(\tau_{1:t+s}\). The expected paid loss can be identified/constructed using chain-ladder approach. Specifically, we use the following relationships introduced in the numerical implementation section:

\[
\mathbb{E}\left[ L_{1}^{(n,t)}\right] = p^{(n,t)} \cdot f_1 \\
\mathbb{E}\left[ \sum_{m=1}^{s} L_{m}^{(n,t+m-1)} | \sum_{m=1}^{s-1} L_{m}^{(n,t+m-1)} \right] = \sum_{m=1}^{s-1} L_{m}^{(n,t+m-1)} \cdot f_{m+1}, \ m = 1, 2, \ldots, d_n - 1,
\]

where \(f_s\)'s are the chain-ladder factor estimated using loss triangles in the last 10 years. Thus, in addition to the chain-ladder factors, we need the premium for each line and we use net premium written \(P_{n,i}\) The interest rate discount factor \(\beta\), though being flat in our theoretical model, can be identified using FRED yield curve of a given year, with \(m\)-year yield denoted by \(r_m\). We denote expected future loss (adjusted for interest rate discount and without adjustment for term structure of reserve) for company \(i\) as \(EL_i\) with definition:

\[
EL_i = \sum_{n=1}^{N} \left( (1 + r_1)^{-1} P_{n,i} \cdot f_1 + \sum_{m=2}^{d_n} \left( 1 + r_m \right)^{-m} P_{n,i} \cdot f_1 \cdots \cdot f_{m-1} \cdot (f_m - 1) \right).
\]

We use the following function from Nelson and Siegel (1987) to start identifying the term structure component.

\[
g(m; \beta_0, \beta_1, \beta_2, \tau) = \left( \beta_0 + \beta_1 \left( 1 - \frac{m}{m + \tau} \right) + \beta_2 \left( 1 - \frac{(m + \tau) e^{-m \tau}}{m} \right) \right), \ m \in (0, \infty).
\]

Even though more complex modelings of yield curve have developed, the above function pro-
vides a flexible fit for any yield curve while adds minimum complexity to the overall modeling. $\beta_0$ describes the long-term value of $f$ (as $m$ goes to infinity). $\beta_0 + \beta_1$ describes the short-term value of $f$ (as $m$ goes to zero). $\beta_2$ gives $f$ function a hump or S shape. $\tau$ can be understood as a tuning parameter that determines how quickly the curve decays from the short-term value to its long-term value.

To help with our estimation goal, we make three assumptions on the term structure. First, the term structure is monotonic, i.e. $\beta_2 = 0$. In this case, a positive $\beta_1$ results in a decreasing yield curve and a negative $\beta_1$ results in an increasing one. Second, the term structure approaches zero in infinite future, or $\lim_{m \to \infty} g(m; \beta_0, \beta_1, \beta_2, \tau) = \beta_0 = 0$. The intuition is that companies are effectively risk neutral about the losses in infinite future and thus evaluate them at actuarial value. Also intuitively from our theoretical results, a firms capitalization level would approach optimality as firm always choose optimal level of exposure and external capital raising in the infinite future, thus $V_1$ will approach zero when time goes to infinity. Third, $\beta_1$ depends on how well the company is capitalized. More precisely, $\beta_1 = b_0 + b_1 \cdot X_i$, where $X_i$ is a variable that assesses the capitalization of a company. Along with our assumptions, we can write down the term structure function as follows:

$$g_i(m) = (b_0 + b_1 X_i) \left( \frac{1 - e^{-\frac{m}{\tau}}}{m} \right),$$

and together with expected paid loss, we have the identification for valuation of total loss reserve:

$$EL_i = b_0 \sum_{n=1}^{N} \sum_{m=1}^{d_n} (1 + r_1)^{-1} \left( \frac{1 - e^{-\frac{m}{\tau}}}{m} \right) \cdot P_{n,i} \cdot f_1 + \sum_{m=2}^{d_n} (1 + r_m)^{-m} \left( \frac{1 - e^{-\frac{m}{\tau}}}{m} \right) \cdot P_{n,i} \cdot f_1 \cdot \ldots \cdot f_{m-1} \cdot (f_{m-1})$$

$$+ b_1 X_i \sum_{n=1}^{N} \sum_{m=1}^{d_n} (1 + r_1)^{-1} \left( \frac{1 - e^{-\frac{m}{\tau}}}{m} \right) \cdot P_{n,i} \cdot f_1 + \sum_{m=2}^{d_n} (1 + r_m)^{-m} \left( \frac{1 - e^{-\frac{m}{\tau}}}{m} \right) \cdot P_{n,i} \cdot f_1 \cdot \ldots \cdot f_{m-1} \cdot (f_{m-1})$$

$VEL_i$ means the valuation of loss reserve for company $i$. Now Equation 9 looks like:

$$P_i - \alpha_i = EL_i + b_0 \cdot VEL_i + b_1 \cdot X_i \times VEL_i + c \cdot SURP_i,$$

where $\alpha_i$ correspond to the addition amount over/under total premium on the left-hand side of Equation 9. We write $\alpha_i = \alpha + \epsilon_i$, with $\mathbb{E}[\epsilon_i] = 0$. We define $EP_i = P_i - EL_i$ as excess premium over expected losses. Rearrange the equation above, we propose the following model for estimation:

$$EP_i = \alpha + b_0 \cdot VEL_i + b_1 \cdot X_i \times VEL_i + c \cdot SURP_i + \epsilon_i \quad (10)$$

The estimation of the model takes two steps. We notice that there is a hidden parameter $\tau$. We follow the empirical approach in Nelson and Siegel (1987) and first find a grid of $\tau$ values. Conditional on $\tau$, the model essentially becomes a linear model and can be estimated using OLS approach. We analyze the model for each $\tau$ on a grid and find the “best” fit with the least residual standard deviation.

We use data obtained from combined annual statements of P&C insurance companies at group level in the U.S. and across multiple years. From each company, we can obtain net premium written and paid loss triangles (Schedule P part 3) for each business line. We have two candidates for $X_i$:...
Table 3: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>2006</th>
<th>2011</th>
<th>2017</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Mean</td>
</tr>
<tr>
<td>$P_i$</td>
<td>454,120</td>
<td>2,448,032</td>
<td>421,235</td>
</tr>
<tr>
<td>$EL_i$</td>
<td>205,130</td>
<td>1,180,483</td>
<td>184,314</td>
</tr>
<tr>
<td>$EP_i$</td>
<td>248,990</td>
<td>1,291,260</td>
<td>236,921</td>
</tr>
<tr>
<td>$SURP_i$</td>
<td>473,793</td>
<td>2,866,170</td>
<td>580,728</td>
</tr>
<tr>
<td>$SA_i$</td>
<td>0.4450</td>
<td>0.1988</td>
<td>0.4638</td>
</tr>
</tbody>
</table>

No. Obs. | 895 | 955 | 900

surplus to asset ratio $SA_i$ and logged leverage ratio $LR_i = \log\left(\frac{1 - SA_i}{SA_i}\right)$. We obtain $SURP_i$, $SA_i$ and $LR_i$ for year 2006, 2011 and 2017. We use 10 years of triangle to compute $EL_i$ and $VEL_i$. For example, we obtain triangles from year 1996–2005 to estimate chain-ladder factors for expected losses in year 2006.

The estimation results are listed in Table 4. We use $\tau$ values from 0.1 to 2 with 0.01 increment and $\tau^*$ is one that results in the least residual standard error. For both candidates of $X_i$, the estimated unit capital cost $c$ is about the same, with about 8% in 2006, 13% in 2011, and 7.5% in 2017. The estimate for $b_0$ and $b_1$ are all significant. Our estimation is robust to additional regressors as shown in Table 5.

We can use the estimates to back out term structure function $g_i(m)$. Figure 12 and 13 shows the term structure in three years for two different companies. The first company has lower capitalization level than the second one. In year 2006 and 2017, both companies has downward sloping term structure, but Company 1 price future expected losses at a much higher level. In year 2011, Company 2 has a upward sloping curve, which is not usually seen and only happens when the company has a very high level of capitalization. Company 2 price the future losses below actuarial value and the premium mostly recoup the high capital costs. Again, the empirical findings are in line with our theoretical results presented in Section 2.5.

![Figure 12: Term structure of loss reserves for representative company with normal capitalization](image-url)
4 Conclusion

In this paper, we set out to explore how a financial institution evaluates its future cash flows as a term structure of capital costs in a P&C insurance company setting. We find both theoretical and empirical implication that a well-capitalized insurer discount its future claim payments into premium and has a upward sloping term structure, while other insurers with lower capitalization level mark up their future claim payments and have a downward sloping term structure.

The model presented in this paper takes into account the loss structure of a P&C insurance company, which is a novel feature relative to the previous literature. The general model is very flexible and can be applied to insurance companies that have both short-tailed and long-tailed business lines. The implementation of the model with two lines and two development years shows previous loss exposure and loss realization significantly affect the company’s optimal policies. We find that long-tailed lines are employed as short-term sources of financing, an insight that considerably changes the characteristics and optimal policies relative to short-tailed lines.

Various extensions are possible. First, for tractability, we adopt a chain-ladder method with normal distributions. In the actuarial literature, there are more advanced models for estimating and forecasting loss triangles that may be considered. Second, we begin with the assumption that all
the assets are invested at a fixed interest rate. Adding securities markets to the general model as well as other financing options or reinsurance would further bridge the gap between model and reality.

References


Appendix

A Technical Appendix

A.1 Proofs of the Lemmas and Propositions

Lemma 2.1 The optimization problem (4) can be equivalently represented as a maximization of the present value of future dividends:

$$\max_c \mathbb{E} \left[ \sum_{t \leq t^*} -\beta^{t-1} B^{(t)} - a^{(0)} \right],$$

where $t^*$ is the time such that $a^{(1)} \geq 0, a^{(2)} \geq 0, \ldots, a^{(t^*-1)} \geq 0, a^{(t^*)} < 0$.

Proof. The capital motion equation (1) can be rewritten into:

$$\beta^t a^{(t)} - \beta^{t-1} a^{(t-1)} - \beta^{t-1} B^{(t)} = \beta^t \left[ (1 + r)p^{(t)} - I^{(t)} - (1 + r)(\tau a^{(t-1)} + c(B^{(t)})) \right]$$

We can rewrite the objective function in (4) as the following:

$$\sum_{t=1}^{\infty} \mathbb{E} \left[ \mathbb{1}_{\{a^{(1)} \geq \ldots, a^{(t)} \geq 0\}} \beta^t \left\{ (1 + r)p^{(t)} - I^{(t)} - \beta^t (1 + r)(\tau a^{(t-1)} + c(B^{(t)})) \right\} \right]$$

$$- \mathbb{1}_{\{a^{(1)} \geq \ldots, a^{(t)} < 0\}} (1 + r)(a^{(t-1)} + B^{(t)})$$

$$= \mathbb{E} \left[ \sum_{t \leq t^*} \beta^t a^{(t)} - \beta^{t-1} a^{(t-1)} - \beta^{t-1} B^{(t)} - \beta^{t^*-1} a^{(t^*-1)} + B^{(t^*)} \right]$$

$$= \mathbb{E} \left[ \sum_{t \leq t^*} -\beta^{t-1} B^{(t)} + \beta^{t^*-1} a^{(t^*-1)} - a^{(0)} - \beta^{t^*-1} a^{(t^*-1)} \right]$$

$$= \sum_{t \leq t^*} \mathbb{E} \left[ -\beta^{t-1} B^{(t)} - a^{(0)} \right]$$

Proposition 2.1 (Bellman Equation). The Bellman equation for problem (4) reads:

$$V(a^{(t-1)}, Q^{(t-1), \Delta^{(t-1)}})$$

$$= \max_{q^{(t)}, p^{(t), B^{(t)}}} \mathbb{E} \left[ \mathbb{1}_{\{I^{(t)} \leq S^{(t)}\}} \left( p^{(t)} - \beta I^{(t)} - \tau a^{(t-1)} - c(B^{(t)}) + \beta V(a^{(t)}, Q^{(t), \Delta^{(t)}}) \right) \right]$$

$$- \mathbb{1}_{\{I^{(t)} > S^{(t)}\}} (a^{(t-1)} + B^{(t)}) \mid \Delta^{(t-1)}$$

subject to (1), (2), (6)

Proof. Since our objective function in (4) is bounded from above, following Bertsekas (1995), the infinite horizon optimization problem (4) subject to (1) is exactly resulting in the Bellman equation (2.1). 

$\square$
Proposition 2.2 (Term Structure Equation). We have the marginal cost of risks presented in the form of term structure of reserve and capital costs:

\[
\begin{align*}
E \left[ \sum_{j=1}^{d_n} \beta^j I_j^{(n,t+j-1) \mid \Delta^{(\cdot,t-1)}} \right] \ast \pi^{(n)} \left( 1 + \sum_{i=1}^{N} \frac{\pi_2^{(i)}}{\pi^{(n)}} E \left[ R^{(i,t) \mid \Delta^{(\cdot,t-1)}} \right] \right) \\
= \sum_{s=0}^{d_n} \beta^{s+1} E \left[ 1 \{ I^{(\cdot,t)} \leq S^{(\cdot,t)}, \ldots, I^{(\cdot,t+s)} \leq S^{(\cdot,t+s)} \} \sum_{s=1}^{n} \left( 1 + V_1(0^{(t+s)}, Q^{(s,t+s)}, \Delta^{(\cdot,t+s)}) \right) (1 - c'(B^{(t)})) \right] \triangle^{(\cdot,t+1)} \\
+ \beta \frac{\partial p(I^{(t)})}{\partial q(n,t)} \left( \sum_{i=1}^{N} E \left[ R^{(i,t) \mid \Delta^{(\cdot,t-1)}} \right] \ast \pi_1^{(i) \frac{\partial \phi}{\partial S^{(t)}}} \right)
\end{align*}
\]

(11)

Proof. The Bellman equation reads:

\[
V(a^{(t-1)}, Q^{(\cdot,t-1)}, \Delta^{(\cdot,t-1)}) = \max_{q^{(\cdot,t)}, \phi^{(\cdot,t)}, B^{(t)}} \left[ 1 \{ I^{(\cdot,t)} \leq S^{(\cdot,t)} \} \left( p^{(t)} - \beta I^{(t)} - \tau a^{(t-1)} - c(B^{(t)}) + \beta V(a^{(t)}, Q^{(\cdot,t)}, \Delta^{(\cdot,t)}) \right) \right] \\
- 1 \{ I^{(\cdot,t)} > S^{(\cdot,t)} \} \left( a^{(t-1)} + B^{(t)} \right) \triangle^{(\cdot,t-1)}
\]

where,

\[
\begin{align*}
I^{(t)} &= \sum_{n=1}^{N} \left( q^{(n,t)} \ast L_1^{(n,t)} + Q^{(n,t-1)} \ast L^{(n,t)} \right) \\
S^{(t)} &= \left( a^{(t-1)}(1 - \tau) + B^{(t)} - c(B^{(t)}) + p^{(t)} \right) (1 + r) \\
R^{(n,t)} &= \sum_{j=1}^{d_n} \beta^{j-1} q^{(n,t)} L_j^{(n,t+j-1)}
\end{align*}
\]

subject to:

\[
\begin{align*}
a^{(t)} &= S^{(t)} - I^{(t)} \\
p^{(n,t)} &= E \left[ \beta R^{(n,t) \mid \Delta^{(\cdot,t-1)}} \right] \ast \pi^{(n)}(\phi, \theta), \quad n = 1, 2, \ldots, N
\end{align*}
\]

The premium function is the product of conditional expected present value of \( R^{(n,t)} \), or future losses of the accident year \( t \), and a markup function \( \pi^{(n)} \). The markup function consists of two arguments: a risk metric \( \phi \) and company size. We assume the risk metric is scale invariant, or \( \phi(w I^{(t)}), \phi(S^{(t)}) = \phi(I^{(t)}, S^{(t)}) \), \( w > 0 \). Denote \( \frac{\partial \pi^{(n)}(x,y)}{\partial x} = \pi_1^{(n)} \) and \( \frac{\partial \pi^{(n)}(x,y)}{\partial y} = \pi_2^{(n)} \).

In addition, we denote the gradients of the value function

\[
V_1(a, Q^{(\cdot)}, \Delta^{(\cdot)}) = \lim_{\delta \to 0} \frac{V(a + \delta, Q^{(\cdot)}, \Delta^{(\cdot)}) - V(a, Q^{(\cdot)}, \Delta^{(\cdot)})}{\delta}
\]

For \( s = 1, 2, \ldots, d_n - 1 \)

\[
V_{2,s}^{(n)}(a, Q^{(\cdot)}, \Delta^{(\cdot)}) = \lim_{\delta \to 0} \frac{V(a, Q^{(1:n-1)}, \ldots, q^{(n)} + \delta, \ldots, Q^{(n+1:N)}, \Delta^{(\cdot)}) - V(a, Q^{(1:n-1)}, \ldots, q^{(n)}, \ldots, Q^{(n+1:N)}, \Delta^{(\cdot)})}{\delta}
\]

\[
V_{2,2}^{(n)}(a, Q^{(\cdot)}, \Delta^{(\cdot)}) = \lim_{\delta \to 0} \frac{V(a, Q^{(1:n-1)}, \ldots, q^{(n)} + \delta, \ldots, Q^{(n+1:N)}, \Delta^{(\cdot)}) - V(a, Q^{(1:n-1)}, \ldots, q^{(n)}, \ldots, Q^{(n+1:N)}, \Delta^{(\cdot)})}{\delta}
\]
The Lagrangian writes:

\[ \mathcal{L}^{(t)} = \mathbb{E} \left[ \mathbf{1}_{\{I^{(t)} \leq S^{(t)}\}} \left( p^{(t)} - \beta I^{(t)} - \gamma a^{(t-1)} - c(\Phi^{(t)}) + \beta V(a^{(t)}, Q^{(c,t)}, \Delta^{(c,t)}) \right) - \mathbf{1}_{\{I^{(t)} > S^{(t)}\}} (a^{(t-1)} + B^{(t)}) \right] \]

\[ - \sum_{i=1}^{N} \lambda^{(i,t)} \left( p^{(i,t)} - \mathbb{E} \left[ \beta R^{(i,t)} \left| \Delta^{(c,t-1)} \right] \right) * \pi^{(i)}(\phi, \theta) \]

Take first order conditions:

\[ \frac{\partial \mathcal{L}^{(t)}}{\partial q^{(n,t)}} = \mathbb{E} \left[ \mathbf{1}_{\{I^{(t)} \leq S^{(t)}\}} \left( -\beta L_1^{(n,t)} + \beta \left( -L_1^{(n,t)} V_1(a^{(t)}, Q^{(c,t)}, \Delta^{(c,t)}) + V_2^{(n)}(a^{(t)}, Q^{(c,t)}, \Delta^{(c,t)}) \right) \right) \right] \Delta^{(c,t-1)} \]

\[ + \lambda^{(n,t)} \left( \sum_{j=1}^{d_n} \beta^j L_j^{(n,t+j-1)} \left| \Delta^{(c,t-1)} \right] \right) * \pi^{(n)}(\phi(I^{(t)}), S^{(t)}), \mathbb{E}_{t-1} \left[ R^{(t)} \left| \Delta^{(c,t-1)} \right] \right) \]

\[ + \sum_{i=1}^{N} \lambda^{(i,t)} \mathbb{E} \left[ \beta R^{(i,t)} \left| \Delta^{(c,t-1)} \right] \right) * \left( \pi^{(i)}(\phi) \frac{\partial \phi}{\partial q^{(n,t)}} \right) + \pi^{(i)}(\phi) \sum_{j=1}^{d_n} \beta^j L_j^{(n,t+j-1)} \left| \Delta^{(c,t-1)} \right] \right) = 0 \quad (12) \]

\[ \frac{\partial \mathcal{L}^{(t)}}{\partial p^{(n,t)}} = \mathbb{E} \left[ \mathbf{1}_{\{I^{(t)} \leq S^{(t)}\}} \left( 1 + V_1(a^{(t)}, Q^{(c,t)}, \Delta^{(c,t)}) \right) \right] \Delta^{(c,t-1)} \]

\[ - \lambda^{(n,t)} + \sum_{i=1}^{N} \lambda^{(i,t)} \mathbb{E} \left[ \beta R^{(i,t)} \left| \Delta^{(c,t-1)} \right] \right) * \pi^{(i)}(\phi) \frac{\partial \phi}{\partial p^{(n,t)}} = 0 \quad (13) \]

\[ \frac{\partial \mathcal{L}^{(t)}}{\partial B^{(t)}} = \mathbb{E} \left[ \mathbf{1}_{\{I^{(t)} \leq S^{(t)}\}} \left( -c'(B^{(t)}) + (1 - c'(B^{(t)})) V_1(a^{(t)}, Q^{(c,t)}, \Delta^{(c,t)}) \right) - \mathbf{1}_{\{I^{(t)} > S^{(t)}\}} \right] \Delta^{(c,t-1)} \]

\[ + \sum_{i=1}^{N} \lambda^{(i,t)} \mathbb{E} \left[ \beta R^{(i,t)} \left| \Delta^{(c,t-1)} \right] \right) * \pi^{(i)}(\phi) \frac{\partial \phi}{\partial B^{(t)}} = 0 \quad (14) \]

The envelope theorem suggests that

\[ V_{2,1}^{(n)}(a^{(t)}, Q^{(c,t)}, \Delta^{(c,t)}) = -\beta \mathbb{E} \left[ \mathbf{1}_{\{I^{(t+1)} \leq S^{(t+1)}\}} L_2^{(n,t+1)} \left( 1 + V_1(a^{(t+1)}, Q^{(c,t+1)}, \Delta^{(c,t+1)}) \right) \right] \]

\[ + \sum_{i=1}^{N} \lambda^{(i,t+1)} \mathbb{E} \left[ \beta R^{(i,t+1)} \left| \Delta^{(c,t+1)} \right] \right) * \pi^{(i)}(\phi) \frac{\partial \phi}{\partial p^{(n,t)}} = 0 \quad (15) \]

\[ V_{2,2}^{(n)}(a^{(t)}, Q^{(c,t)}, \Delta^{(c,t)}) = -\beta \mathbb{E} \left[ \mathbf{1}_{\{I^{(t)} \leq S^{(t)}\}} L_2^{(n,t+1)} \left( 1 + V_1(a^{(t+1)}, Q^{(c,t+1)}, \Delta^{(c,t+1)}) \right) \right] \]

\[ + \sum_{i=1}^{N} \lambda^{(i,t+2)} \mathbb{E} \left[ \beta R^{(i,t+2)} \left| \Delta^{(c,t+1)} \right] \right) * \pi^{(i)}(\phi) \frac{\partial \phi}{\partial B^{(t)}} = 0 \]
Now, since

\[ V^{(n)}_{2,d_n-1}(a^{(t+d_n-2)}, Q^{(t+d_n-2)}, \Delta^{(t+d_n-2)}) \]

\[ = -\beta \mathbb{E} \left[ I_{(1 \leq z \leq 1+t) \leq S^{*} \leq (1+t)} L^{(n,t+d_n-1)}_{2} \left( 1 + V_1(a^{*}, Q^{*}, \Delta^{*}) \right) \right] + V^{(n)}_{2,d_n} \left( S^{*}, \Delta^{*} \right) \]

From equations (13) and (14), we have

\[ \frac{\partial \phi}{\partial B(t)} = \frac{\partial \phi}{\partial S(t)} = \frac{\partial \phi}{\partial B(t)} = \frac{\partial \phi}{\partial S(t)} = \frac{\partial \phi}{\partial S(t)} = \frac{\partial \phi}{\partial S(t)} = \frac{\partial \phi}{\partial S(t)} \]

From equations (13) and (14), we have

\[ \chi^{(n,t)} = \frac{1}{1 - c'(B(t))}, \quad \forall n = 1, 2, \ldots, N, \quad t = 1, 2, \ldots \]

and

\[ \sum_{i=1}^{N} \frac{1}{1 - c'(B(t))} \mathbb{E} \left[ R^{(i,t)} | \Delta^{(t-1)} \right] \mathbb{P}^{(i,t)} \left( \frac{\partial \phi}{\partial S(t)} \right) = \frac{1}{1 - c'(B(t))} \mathbb{E} \left[ I_{(1 \leq z \leq 1+t) \leq S^{*}} \left( 1 + V_1(a^{*}, Q^{*}, \Delta^{*}) \right) \right] \]

The scale invariance property of \( \phi \) yields the following (cf. Bauer and Zanjani, 2018):

\[ 0 = \frac{\partial}{\partial \omega} \phi(w I(t), w S(t)) = S(t) \frac{\partial}{\partial S(t)} \phi(w I(t), w S(t)) + I(t) \frac{\partial}{\partial I(t)} \phi(w I(t), w S(t)) \]

Define \( \rho \) as the risk measure associated with the risk metric \( \phi \), with adding-up property:

\[ \sum_{i=1}^{N} \sum_{j=0}^{d_n-1} \frac{\partial \phi}{\partial q^{(i,t-j)}} = \sum_{i=1}^{N} \sum_{j=0}^{d_n-1} \frac{\partial \phi}{\partial q^{(i,t-j)}} = S(t) \]
Hence for every $i$, $j$ and $t$, we have:

$$\frac{\partial \phi}{\partial q^{(n,t-j)}} = -\frac{\partial \phi}{\partial S^{(t)}} \ast \frac{\partial p(I^{(t)})}{\partial q^{(n,t-j)}} \tag{19}$$

With equations (16), (17) and (19), (12) becomes:

$$\frac{\partial L^{(t)}}{\partial q^{(n,t)}} = -\beta E\left[ I_{\{I^{(t)} \leq S^{(t)}\}} \left(L^{(n,t)}_1 + V_1(a^{(t)}, Q^{(c,t)}, \Delta^{(c,t,s)})\right) | \Delta^{(c,t-1)} \right]$$

$$- \sum_{s=1}^{d_n-1} \beta^{s+1} E\left[ I_{\{I^{(t)} \leq S^{(t)}, \ldots, I^{(t+s)} \leq S^{(t+s)}\}} L^{(n,t+s)}_s + V_1(a^{(t+s)}, Q^{(c,t+s)}, \Delta^{(c,t+s)}) | \Delta^{(c,t-1)} \right]$$

$$+ \frac{1}{1-c'(B(t))} \sum_{j=1}^{d_n} \beta^j L^{(n,t+j-1)} \ast \Delta^{(c,t-1)} \ast \left( \pi^{(n)} + \sum_{i=1}^{N} \pi^{(i)} \ast \frac{\partial \phi}{\partial q^{(n,t)}} \right)$$

Rearrange and obtain the term structure equation (7).

\[ \square \]

### A.2 2L2DY Loss Distribution Assumptions

In loss triangle depicted by Figure 3, $L_1^{(n)}$ are losses paid in the previous accident year $t-1$. $L_1^{(1)}$ is a realization included in the Markov structure, and therefore is a state variable in the optimal control problem. Line 2 has no development year, so $L_1^{(2)}$ is irrelevant in the optimal control problem. $L_1^{(1)}, L_1^{(2)},$ and $L_2^{(1)}$ are the losses to be paid in the current period $t$, with $L_1^{(1)}$ and $L_1^{(2)}$ being paid losses for accident year $t$ and $L_2^{(1)}$ being the second development year paid loss for the previous accident year $t-1$ of line 1. Therefore, $L_1^{(1)}, L_1^{(2)},$ and $L_2^{(1)}$ are the stochastic disturbances and have distributions subject to probability measures $p(dL_1^{(1)} | L_1^{(1)}), p(dL_1^{(2)} | L_1^{(2)})$ and $p(dL_2^{(1)} | L_1^{(1)})$. Meanwhile, $L_2^{(1)}$ is the paid loss in the next period $t+1$, developed from $L_1^{(1)}$ and therefore related to the current period’s premium in Line 1, $L_2^{(1)}$ is also the stochastic disturbance subject to probability measures $p(dL_2^{(1)} | L_1^{(1)})$. We make the following assumptions for the properties of the probability measures:

Following **Chain-Ladder**, we assume:

- \( \mathbb{E}(L_2^{(1)} | L_1^{(1)}) = (f-1)L_1^{(1)} \),
- \( \mathbb{V}(L_2^{(1)} | L_1^{(1)}) = \sigma^2 L_1^{(1)} \).

We assume **conditional normality**:

- \( L_1^{(n)} | L_1^{(1)} \sim N(\mu_1^{(n)}, (\sigma_1^{(n)})^2) \quad n = 1, 2 \),
- \( L_2^{(1)} | L_1^{(1)} \sim N((f-1)L_1^{(1)}, \sigma^2 L_1^{(1)}) \),
- \( L_2^{(1)} | L_1^{(1)} \sim N((f-1)L_1^{(1)}, \sigma^2 L_1^{(1)}) \).
We assume a **linear correlation** between lines:

\[ \text{corr}(L_1^{(1)}, L_1^{(2)}) = \text{corr}(L_1^{(1)}, L_1^{(2)}) = \rho. \]

The chain-ladder property follows Mack (1993), whose model assumes \( M_{t,j-1}^{(n)} = C_{t,j-1}^{(n)}. \) That is, the cumulative paid loss in each accident year is a Markov chain, with the ultimate development year being the time horizon of the chain. \( f \) and \( \sigma^2 \) are respectively the chain-ladder factor and its variance factor. Both factors can be estimated using a loss triangle. More precisely referring to Figure 2, using the past loss realization from accident year 1 to \( t-2 \), we can estimate the chain ladder factors following:

\[
\hat{f} = \frac{\sum_{k=1}^{t-2} L_1^{(1,k)} + L_2^{(1,k+1)}}{\sum_{k=1}^{t-2} L_1^{(1,k)}},
\]

\[
\hat{\sigma}^2 = \frac{1}{t-3} \sum_{k=1}^{t-2} \left( L_1^{(1,k)} + L_2^{(1,k+1)} \right) \left[ \frac{L_1^{(1,k)} + L_2^{(1,k+1)}}{L_1^{(1,k)}} - \hat{f} \right]^2.
\]

As is conventional in this context, the indemnity is assumed to be proportional to the current-period exposures \( q^{(n)} \) and last-period exposures \( q^{(n)}. \) Hence, the indemnity random variable is specified as

\[
I = q^{(1)} L_1^{(1)} + q^{(2)} L_1^{(2)} + q^{(1)} L_2^{(1)}. \]

This linearity assumption entails that the marginal claim distribution is fixed, so that the loss distribution is homogeneous. In addition to linearity of indemnity, the conditional normality assumption ensures that \( I \) is also normal, with conditional mean and variance:

\[
\mu_I = \mathbb{E}(I | L_1^{(1)}) = q^{(1)} \mu_1^{(1)} + q^{(2)} \mu_1^{(2)} + (f - 1) q^{(1)} L_1^{(1)} ,
\]

\[
\sigma_I^2 = \mathbb{V}(I | L_1^{(1)}) = (q^{(1)})^2 (\sigma_1^{(1)})^2 + (q^{(2)})^2 (\sigma_1^{(2)})^2 + (q^{(1)})^2 \sigma_1^2 L_1^{(1)} + 2 q^{(1)} q^{(2)} \rho \sigma_1^{(1)} \sigma_1^{(2)} .
\]

Again, this is in line with typical assumptions, and generalizations are possible.

We estimate model parameters for the accident year loss from \( k = 1 \) to \( t - 1 \), using a simple regression with linear time trend:

\[
L_1^{(n,k)} = \xi_0^{(n)} + \xi_1^{(n)} k + \epsilon_k^{(n,k)}, \quad \epsilon_k^{(n,k)} \sim \mathcal{N}(0, (\sigma_k^{(n)})^2),
\]

so that:

\[
\hat{\mu}_1^{(n)} = \hat{\xi}_0^{(n)} + \hat{\xi}_1^{(n)} t ,
\]

\[
(\hat{\sigma}_1^{(n)})^2 = \sum_k \frac{\left( \hat{\epsilon}_1^{(n,k)} \right)^2}{t-3},
\]

\[
\hat{\rho} = \text{corr}(\hat{\epsilon}_1^{(1,k)}, \hat{\epsilon}_1^{(2,k)}),
\]

although the Bellman equation assumes identical distributions going forward.
A.3 Numerical Solution of 2L2DY Model

To solve the Bellman equation, we rely on numerical methods. By our premium function assumption, \( q^{(1)} \), \( q^{(2)} \), and \( B \) endogenously determine the sum of premiums \( P \), and therefore these are the only choice variables. Note that the expectations of Equation (8) entail functions of \( L^{(1)}_1 \) and \( I \), which renders the problem two-dimensional. To solve it, we use the numerical integration method from Tanskanen and Lukkarinen (2003).

First, let \( X = L^{(1)}_1 \) and \( Y = I | L^{(1)}_1 \). Note that \( X \) and \( Y \), both univariate normal, can be represented as a bivariate normal distribution:

\[
\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N} \left( \begin{pmatrix} \mu^{(1)}_I \\ \mu_I \end{pmatrix}, \begin{pmatrix} (\sigma^{(1)}_I)^2 & \rho_{x,y} \sigma^{(1)}_I \sigma_I \\ \rho_{x,y} \sigma^{(1)}_I \sigma_I & \sigma^2_I \end{pmatrix} \right),
\]

where \( \rho_{x,y} = \frac{q^{(1)} \sigma^{(1)}_I + q^{(2)} \sigma^{(2)}_I}{\sigma_I} \). Hence, the conditional distribution of \( Y \) given \( X \) is also univariate normal with mean and variance

\[
\begin{align*}
\mu_{y|x} &= \mu_I + \rho_{x,y} \sigma_I \frac{x - \mu^{(1)}_I}{\sigma^{(1)}_I}, \\
\sigma_{y|x}^2 &= \sigma^2_I (1 - \rho^2_{x,y}),
\end{align*}
\]

respectively.

Let \( f(x, y) \) be the density of the bivariate normal of \( X \) and \( Y \), \( f(y|x) \) be the conditional density of \( Y \) given \( X \), and \( f(x) \) be the marginal density of \( X \). Because of the nature of the value function is unknown, we need to use the value iteration method to solve the Bellman equation (8) on a discretized state-space. For 2L2DY model, there are three state variables: capital \( a \), last-period exposures on long-tailed line 1 \( q^{(1)} \), and last-period loss realizations on long-tailed line 1 \( L^{(1)}_1 \).

Here are the steps of solving the Bellman equation using value iteration.

1. Pick grids for \( a = (a_1, a_2, \ldots, a_s) \), \( q^{(1)} = (q_1, q_2, \ldots, q_n) \), and \( L^{(1)}_1 = (x_1, x_2, \ldots, x_p) \). Set \( V_0 = v_0(a, q^{(1)}, L^{(1)}_1) \), where \( v_0 \) is an arbitrary function.
2. Solve the optimization problem on the right hand side of the Bellman equation and get optimized state variables \( c^* \) and yield policy function \( c = u_1((a, q^{(1)}, L^{(1)}_1); c^*) \). Then obtain the next value function \( V_1(\{(a, q^{(1)}, L^{(1)}_1); u_1\}) \) until \( V_j \) converges.

We can obtain a simplified Bellman equation for implementation:

\[
\beta \mathbb{E} \left\{ 1_{\{I \leq S\}} S - I \right\} = \int_{-\infty}^{S} \beta (S - y) f(y) \, dy
\]

\[
= \beta S \Phi_I(S) - \beta \int_{-\infty}^{S} y \frac{1}{\sqrt{2\pi} \sigma_I} e^{-\frac{(y-\mu_I)^2}{2\sigma_I^2}} \, dy
\]

\[
= \beta (\Phi_I(S) + \sigma_I \phi_I(S))
\]
And
\[
\mathbb{E}\left\{ 1_{\{t \leq s\}} V(a', L_1^{(1)}) \right\} = \int_{-\infty}^{\infty} \int_{-\infty}^{S} V(a', q^{(1)}, x) f(x, y) \, dy \, dx \\
= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{S} V(a', q^{(1)}, x) f(y|x) \, dy \right) f(x) \, dx,
\]
(22)

where \( \Phi \) and \( \phi \), respectively, are the CDF and PDF of a standard normal distribution.

To solve the inner integral on a grid, we apply the Tanskanen and Lukkarinen (2003) method. First, we interpolate on \( a \). We pick \((l+1)\)-point grids for \( I \geq 0 \), say \((y_0, y_1, \ldots, y_l)\), with \( 0 = y_0 < y_1 < \cdots < y_l = S \), let \( \varphi_i = V(a'(y_i), q^{(1)}, L_1^{(1)}) \).

For \( a'(y_i) \in (a_k, a_{k+1}) \), we approximately have by linear interpolation:
\[
\varphi_i = \frac{a_{k+1} - a'(y_i)}{a_{k+1} - a_k} V(a_k, q^{(1)}, L_1^{(1)}) + \frac{a'(y_i) - a_k}{a_{k+1} - a_k} V(a_{k+1}, q^{(1)}, L_1^{(1)})
\]

If \( a'(y_i) > a_l \), we can extrapolate:
\[
\varphi_i = \frac{a'(y_i) - a_{l-1}}{a_l - a_{l-1}} V(a_l, q^{(1)}, L_1^{(1)}) + \frac{a'(y_i) - a_l}{a_l - a_{l-1}} V(a_{l-1}, q^{(1)}, L_1^{(1)})
\]

The linear interpolation w.r.t. \( Y \) is
\[
V(a', q^{(1)}, L_1^{(1)}) = \sum_{k=0}^{l-1} \left( \varphi_k + \frac{y - y_k}{y_{k+1} - y_k} \left( \varphi_{k+1} - \varphi_k \right) \right) 1_{[y_k, y_{k+1})}(y)
\]

We then break down the integral into sums:
\[
\int_{-\infty}^{S} V(a', q^{(1)}, x) f(y|x) \, dy \\
= \sum_{k=0}^{l-1} \left[ \left( \varphi_k - \frac{y_k(\varphi_{k+1} - \varphi_k)}{y_{k+1} - y_k} \right) \int_{y_k}^{y_{k+1}} f(y|x) \, dy + \left( \frac{\varphi_{k+1} - \varphi_k}{y_{k+1} - y_k} \right) \int_{y_k}^{y_{k+1}} y f(y|x) \, dy \right] \\
= \sum_{k=0}^{l-1} \left\{ \left( \varphi_k - \frac{y_k(\varphi_{k+1} - \varphi_k)}{y_{k+1} - y_k} \right) \left[ \Phi \left( \frac{y_{k+1} - \mu_{y|x}}{\sigma_{y|x}} \right) - \Phi \left( \frac{y_k - \mu_{y|x}}{\sigma_{y|x}} \right) \right] \\
+ \left( \frac{\varphi_{k+1} - \varphi_k}{y_{k+1} - y_k} \right) \left[ \mu_{y|x} \left( \Phi \left( \frac{y_{k+1} - \mu_{y|x}}{\sigma_{y|x}} \right) - \Phi \left( \frac{y_k - \mu_{y|x}}{\sigma_{y|x}} \right) \right) \right] \\
- \sigma_{y|x} \left( \frac{y_{k+1} - \mu_{y|x}}{\sigma_{y|x}} \right) - \frac{\varphi_{k+1} - \varphi_k}{y_{k+1} - y_k} \left( \frac{y_k - \mu_{y|x}}{\sigma_{y|x}} \right) \right\} \\
= h(x)
\]
(23)
Therefore the right-hand side of our Bellman equation can be written as

\[ \beta \left( (S - \mu_I) \Phi \left( \frac{S - \mu_I}{\sigma_I} \right) + \sigma_I \phi \left( \frac{S - \mu_I}{\sigma_I} \right) \right) + \int_{-\infty}^{\infty} \beta h(x) f(x) dx - a - B, \] (**)

which can be solved using a discretized grid of \( L_{1}^{(1)} \).

For our interpolation, we choose a \( m+1 \)-point equally spaced grid on \([\mu_{1}^{(1)} - 5\sigma_{1}^{(1)}, \mu_{1}^{(1)} + 5\sigma_{1}^{(1)}]\), say \((x_0, x_1, \ldots, x_m)\). In our case the grid size is 26. Use the trapezoidal rule to break the integral down into sums, say \( F(x) = \beta(g(x) + h(x)) f(x) \), then the integral becomes:

\[ \int_{-\infty}^{\infty} F(x) \, dx = \frac{x_m - x_0}{2m} (F(x_0) + 2F(x_1) + \cdots + 2F(x_{m-1}) + F(x_m)) \]

Hence we successfully convert double integrals into sums and therefore significantly reduce the computation time without compromising the accuracy. The value iteration is implemented and run in Julia. The optimization is executed using Julia’s NLopt package. The value function is defined on a 21 x 21 x 3 discretized grid (with 21 grid points on \( a \) and \( q^{(1)} \)). We ran the program for 80 iterations and the value function converges for both value function and choice variables.

The total runtime was about 100,000s, on a Intel I7 dual-core CPU. In this specific numerical task, Julia is six times faster than the popular high-level language such as R and MatLab. In particular, Julia is much faster with loops, which is heavily found in the iterations and numerical integrals. According to Julia language developers, Julia is a high-performance language suitable for dynamic programming and its syntax is easily adapted from R or Matlab. Julia’s high efficiency helps shorten the runtime from one week that would have taken on R, to just under one day.

Compared to previous models without considering DY, which only has one state variable capital, the 2L2DY model suffers from “the curse of dimensionality”. As the general \( nL_jDY \) model has hundreds, million, or even trillion times more states, each iteration of value function would take proportional more time to complete, resulting the value iteration to finish in months or even years. Solving this high-dimensional problem seems infeasible even five years ago, but thanks to the power of modern day computing, it is feasible under proper assumptions. We start with solving 2L2DY, which has the least dimension in the general \( nL_jDY \) model. In the future, we will continue to refine the algorithm and implement parallel computing to further shorten the running time.

B Additional Figures
Figure 14: 2-dimensional representations of external capital raising (small previous shock)

Figure 15: 2-dimensional representations of external capital raising (large previous shock)
Figure 16: Convergence of value function and choice variables