IMPLEMENTATION OF PROPORTIONAL HAZARDS TRANSFORMS IN RATEMAKING

SHAUN WANG

Abstract

This article introduces a relatively new method for calculating risk load in insurance ratemaking: the use of proportional hazards (PH) transforms. This method is easy to understand, simple to use, and supported by theoretical properties as well as economic justification. After an introduction of the PH-transform method, examples show how it can be used in pricing ambiguous risks, excess-of-loss coverages, increased limits, risk portfolios, and reinsurance treaties.

ACKNOWLEDGEMENT

I would like to thank Gary Venter for his inspiration and numerous helpful comments and suggestions. I have also benefited from comments by James Higgins, Israel Krakowski, John Kulik, and the Committee on Review of Papers.

1. INTRODUCTION

Recently, there has been considerable interest in and extensive discussion of risk loads within the Casualty Actuarial Society. These discussions have focused on measures of risk and methods to arrive at a 'reasonable' risk load. Although there are diverse opinions on the appropriate measurement of risk, there is general agreement on the distinction between process risk and parameter risk, and on the importance of parameter risk in ratemaking. (See Finger [5], Miccolis [18], McClenahan [15], Feldblum [4], Philbrick [21], Meyers [16], Robbin [25], and Bault [1].)

Consistent with previous papers, this paper will consider only pure risk-adjusted premiums (the expected loss plus risk load, excluding all expenses and commissions). These pure risk-adjusted premiums are sometimes referred to as premiums in the paper.

Following Venter's [28] advocacy of adjusted distribution methods, Wang [30] proposes using proportional hazards (PH) transforms in the calculation of the risk-adjusted premium. This paper focuses on the practical aspects of implementation of PH-transforms in ratemaking. More specifically, the paper shows how PH-transforms can be used to quantify process risk, parameter risk, and dependency risk. It also discusses economic justification after introducing implementation issues.

To utilize the PH-transform in ratemaking, a probability distribution for claims is needed. A probability distribution can often be estimated from industry claim data or by computer simulations. Even though a probability distribution can be obtained from past claim data, sound and knowledgeable judgements are always required to ensure that the estimated loss distribution is valid for ratemaking.

It is safe to say that no theoretical risk-load formula can claim to be the *right* one, since subjective elements always exist in any practical exercise of ratemaking. However, a good theoretical risk-load formula can assist actuaries and help maintain logical consistency in the ratemaking process. In this respect, it is hoped that the PH-transform method becomes a useful tool for practicing actuaries in insurance ratemaking.

The remainder of this paper is divided into four sections. Section 2 introduces the PH-transform method and applies it to the pricing of a single risk (including excess cover and increased limits ratemaking). Section 3 discusses the use of PH-transforms in pricing risk portfolios and reinsurance treaties. Section 4 discusses two simple mixtures of PH-transforms. The first mixture can yield a minimal rate-on-line, and the second mixture suggests a new measure for the right tail risk. Section 5 briefly reviews the leading economic theories of risk and uncertainty, and their relationship to insurance ratemaking.

2. PROPORTIONAL HAZARDS TRANSFORM

In the pricing of insurance risks, it is common for the actuary to first obtain a best-estimate loss distribution based on all possible information (e.g., empirical data) and/or judgement. The best estimate loss distribution, serving as an anchor, is then transformed into a heavier-tailed distribution, and the mean from the latter is used to price the business, thereby producing a risk load. Venter [28] advocated the adjusted distribution principle and gave a theoretical justification by using a no-arbitrage pricing argument. He observed that the only methods of premium calculation that preserve layer additivity are those that can be generated from transformed distributions, where the premium for any layer is the expected loss for that layer under the transformed distribution. Inspired by Venter's insightful observation, Wang [30] proposed the proportional hazards transform method which is also the topic of this paper.

An insurance risk refers to a non-negative loss random variable X, which can be described by the decumulative distribution function (ddf): $S_X(u) = \Pr\{X > u\}$. An advantage of using the ddf is the unifying treatment of discrete, continuous, and mixed-type distributions. In general, for a risk X, the expected loss can be evaluated directly from its ddf:

$$\mathrm{E}[X] = \int_0^\infty S_X(u)\,du.$$

(A proof of this statement is given in Appendix A.) In practice, the actuary does not know the true underlying loss distribution, but instead may have a best-estimate loss distribution based on available information. The PH-transform is a method for adjusting the best-estimate distribution according to the levels of uncertainty, market competition, and portfolio diversification.

DEFINITION 1 *Given a best-estimate loss distribution* $S_X(u) = \Pr\{X > u\}$, *for some exogenous* ***index*** r $(0 < r \leq 1)$, *the* ***proportional hazards (PH) transform*** *refers to a mapping* $S_Y(u) = [S_X(u)]^r$, *and the* ***PH-mean*** *refers to the expected value under the transformed distribution:*

$$H_r[X] = \int_0^{\infty} [S_X(u)]^r \, du, \qquad (0 < r \leq 1).$$

The PH-mean was introduced by Wang to represent the risk-adjusted premium (the expected loss plus risk load). As we shall see, the PH-mean is quite sensitive to the choice of the index r. It could be infinite for some unlimited loss distributions and choices of r.

EXAMPLE 1 The following three loss distributions,

$$S_U(u) = 1 - u/(2b), \qquad 0 \leq u \leq 2b \qquad \text{(uniform)},$$
$$S_V(u) = e^{-u/b} \qquad \text{(exponential), and}$$
$$S_W(u) = b^2/(b+u)^2 \qquad \text{(Pareto)},$$

have the same expected loss, b. One can easily verify that

$$H_r[U] = \frac{2b}{1+r},$$
$$H_r[V] = \frac{b}{r},$$
$$H_r[W] = \begin{cases} \dfrac{b}{2r-1}, & r > 0.5; \\ \infty, & r \leq 0.5. \end{cases}$$

The PH-mean, interpreted as the risk-adjusted premium, preserves the usually accepted ordering of riskiness based on heaviness of tail (see Table 1). Here it is assumed that the distributions are known to be of the type shown, whereas uncertainty about the type of distribution could contribute further risk.

TABLE 1

SOME VALUES OF THE PH-MEAN $H_r[\cdot]$

	U	V	W
$r_1 = \frac{5}{6}$	$1.09b$	$1.2b$	$1.5b$
$r_2 = \frac{2}{3}$	$1.2b$	$1.5b$	$3.0b$

EXAMPLE 2 When X has a Pareto distribution with parameters (α, λ),

$$S_X(u) = \left(\frac{\lambda}{\lambda + u}\right)^{\alpha}, \qquad \text{and}$$

the PH-transform $S_Y(u)$ also has a Pareto distribution with parameters $(r\alpha, \lambda)$.

When X has a Burr distribution with parameters (α, λ, τ),

$$S_X(u) = \left(\frac{\lambda}{\lambda + u^{\tau}}\right)^{\alpha}, \qquad \text{and}$$

the PH-transform $S_Y(u)$ also has a Burr distribution with parameters $(r\alpha, \lambda, \tau)$.

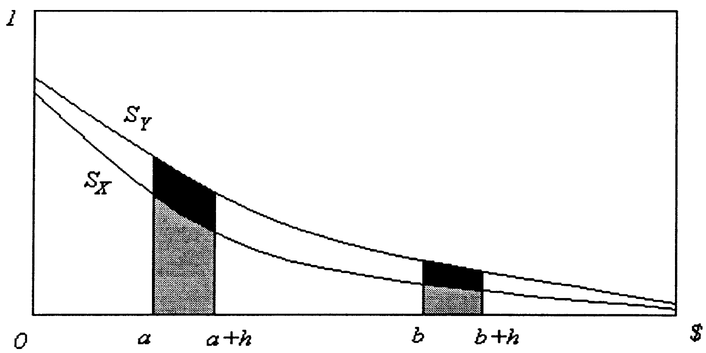
When X has a gamma (or lognormal) distribution, the PH-transform $S_Y(u)$ is no longer a gamma (or lognormal). In such cases, numerical integration may be required to evaluate the PH-mean.

2.1. *Pricing of Ambiguous Risks*

In practice, the underlying loss distribution is seldom known with precision. There are always uncertainties regarding the best-estimate loss distribution. Insufficient data or poor quality data often result in sampling errors. Even if a large amount of high-quality data is available, due to changes in the claim generating mechanisms, past data may not fully predict the future claim distribution. The PH-transform can be adjusted to give a higher risk load when this parameter uncertainty is greater.

FIGURE 1

MARGINS FOR PARAMETER UNCERTAINTY BY PH-TRANSFORMS



As illustrated in Figure 1, the PH-transform, $S_Y(u) = [S_X(u)]^r$, can be viewed as an upper confidence limit for the best-estimate loss distribution $S_X(u)$. A smaller index r yields a wider range between the curves S_Y and S_X. This upper confidence limit interpretation has support in statistical estimation theory (see Appendix B). The index r can be assigned accordingly with respect to the level of confidence in the estimated loss distribution. The more ambiguous the situation is, the lower the value of r that should be used.

EXAMPLE 3 Consider the following experiment conducted by Hogarth and Kunreuther [6]. An actuary is asked to price warranties on the performance of a new line of microcomputers. Suppose that the cost of repair is $100 per unit, and there can be at most one breakdown per period. Also, suppose that the risks of breakdown associated with any two units are independent. The best-estimate of the probability of breakdown has three scenarios:

$$\theta = 0.001, \qquad \theta = 0.01, \qquad \theta = 0.1.$$

The level of confidence regarding the best estimate has two scenarios:

Non-ambiguous: There is little ambiguity regarding the best-estimate loss distribution. Experts all agree with confidence on the chances of a breakdown.

Ambiguous: There is considerable ambiguity regarding the best-estimate loss distribution. Experts disagree and have little confidence in the estimate of the probabilities of a breakdown.

Note that the loss associated with a computer component can only assume two possible values, either zero or \$100. For any fixed $u < 100$, the probability that the loss exceeds u is the same as the probability of being exactly \$100, namely θ. For a fixed $u \geq 100$, it is impossible that the loss exceeds u. Thus, the best-estimate ddf of the insurance loss cost is

$$S_X(u) = \begin{cases} \theta, & 0 < u < 100; \\ 0, & 100 \leq u. \end{cases}$$

The PH-transform with index r yields a risk-adjusted premium of $100\theta^r$.

In both cases a risk load is needed because there is frequency uncertainty, but more load is needed in the ambiguous case. If we choose $r = 0.97$ for the non-ambiguous case, and $r = 0.87$ for the ambiguous case, we get the premium structure shown in Table 2.

For comparison purposes, Table 2 also shows the premium structure using the standard deviation method[1] set to agree with the PH-mean at the 0.01 frequency. Note that for the Bernoulli type of risks in this example, the standard deviation loads vary more by frequency. However, as we shall see in Section 3.2, this

[1]The traditional standard deviation method calculates a risk-adjusted premium by the formula $E[X] + \beta\sigma[X]$, where $\beta \geq 0$ is an exogenous constant.

TABLE 2

THE RATIO OF THE RISK-ADJUSTED PREMIUM TO THE EXPECTED LOSS

PH-Transform Method	$\theta = 0.001$	$\theta = 0.01$	$\theta = 0.1$
Non-ambiguous ($r = 0.97$)	1.23	1.15	1.07
Ambiguous ($r = 0.87$)	2.45	1.82	1.35

Standard Deviation Method	$\theta = 0.001$	$\theta = 0.01$	$\theta = 0.1$
Non-ambiguous ($\beta = 0.01508$)	1.48	1.15	1.05
Ambiguous ($\beta = 0.0824$)	3.60	1.82	1.25

pattern no longer holds for continuous-type risks. The main problem with standard deviation is in its lack of additivity when a risk is divided into sub-layers.

In summary, the PH-transform can be used as a means of provision for estimation errors. The actuary can subsequently set up a table for the index r according to different levels of ambiguity, such as the following:

Ambiguity Level	Index r
Slightly Ambiguous	0.960 – 1.000
Moderately Ambiguous	0.900 – 0.959
Highly Ambiguous	0.800 – 0.899
Extremely Ambiguous	0.500 – 0.799

Note that the premium developed is particularly sensitive to the choice of r, especially for small r, so care should be exercised in its selection.

2.2. *Pricing Excess Layers of a Single Risk*

Since most insurance contracts contain clauses such as a deductible and a maximum limit, it is convenient to use the general language of excess-of-loss layers. A layer $(a, a+h]$ of a risk X is defined by the loss function:

$$X_{(a,a+h]} = \begin{cases} 0, & 0 \le X < a; \\ (X - a), & a \le X < a + h; \\ h, & a + h \le X, \end{cases}$$

where a is the *attachment point (retention)*, and h is the *limit*.

In this subsection, we restrict our discussion to a single risk X (individual or aggregate). For instance, X may represent an underlying risk for facultative reinsurance, or the aggregate loss amount for a risk portfolio being priced. Under this restriction, there will be either no or one claim to a given layer. In other words, the claim frequency to a given layer is Bernoulli. In Section 3 we will discuss the pricing of excess layers of reinsurance treaties where there can be multiple claims to a given layer.

One can verify that the loss variable $X_{(a,a+h]}$ has a ddf of

$$S_{X_{(a,a+h]}}(u) = \begin{cases} S_X(a+u), & 0 \le u < h \\ 0, & h \le u, \end{cases}$$

and that the average loss cost for the layer $(a, a+h]$ is

$$\mathrm{E}[X_{(a,a+h]}] = \int_0^h S_X(a+u)\,du = \int_a^{a+h} S_X(u)\,du.$$

Under the PH-transform $S_Y(u) = [S_X(u)]^r$, the PH-mean for the layer $(a, a+h]$ is

$$\begin{aligned} H_r[X_{(a,a+h]}] &= \int_0^\infty [S_{X_{(a,a+h]}}(u)]^r\,du \\ &= \int_0^h [S_X(a+u)]^r\,du = \int_a^{a+h} [S_X(u)]^r\,du. \end{aligned}$$

In other words, the expected loss and the risk-adjusted premium for the layer $(a, a+h]$ are represented by the areas over the interval $(a, a+h]$ under the curves $S_X(u)$ and $S_Y(u)$, respectively (see Figure 1).

In Wang [30], it is shown that, for $0 < r < 1$, the PH-mean has the following properties:

- Positive loading: $H_r[X_{(a,a+h]}] \geq \mathrm{E}[X_{(a,a+h]}]$.
- Decreasing risk-adjusted premiums:

$$\text{For} \quad a < b, \qquad H_r[X_{(a,a+h]}] \geq H_r[X_{(b,b+h]}].$$

- Increasing relative loading:

$$\text{For} \quad a < b, \qquad \frac{H_r[X_{(a,a+h]}]}{\mathrm{E}[X_{(a,a+h]}]} \leq \frac{H_r[X_{(b,b+h]}]}{\mathrm{E}[X_{(b,b+h]}]}.$$

These properties are consistent with market premium structures (Patrick, [20]; Venter, [28]).

EXAMPLE 4 A single (ground-up) risk has a 10% chance of incurring a claim, and, if a claim occurs, the claim size has a Pareto distribution ($\lambda = 2{,}000$, $\alpha = 1.2$). Putting the Bernoulli frequency and the Pareto severity together, we have a ground-up loss distribution

$$\begin{aligned} S_X(u) &= \Pr\{X > u\} \\ &= \text{Probability of occurrence} \times \Pr\{\text{Loss Size } > u\} \\ &= 0.1 \times \left(\frac{2{,}000}{2{,}000 + u}\right)^{1.2}. \end{aligned}$$

The actuary is asked to price various layers of the (ground-up) risk. Suppose that the actuary infers an index, say $r = 0.92$, from individual risk analysis and market conditions. The actuary may need to consider the risk loads for other contracts with similar characteristics in the insurance and/or financial markets.

TABLE 3

LAYER PREMIUMS USING PH-TRANSFORMS

Layer $X_{(a,b]}$ 1K = $1,000	Expected Loss	$H_{0.92}[X_{(a,b]}]$ ($r = 0.92$)	Percentage Loading	$H_{0.90}[X_{(a,b]}]$ ($r = 0.90$)	Percentage Loading
(0, 1K]	77.89	95.47	22.6%	100.45	29.0%
(5K, 6K]	20.51	27.99	36.5%	30.25	47.5%
(10K, 11K]	11.098	15.91	43.3%	17.41	56.8%
(50K, 51K]	1.982	3.26	64.5%	3.69	86.3%
(100K, 101K]	0.888	1.56	75.4%	1.79	102%
(500K, 501K]	0.132	0.269	104%	0.322	144%
(1,000K, 1,001K]	0.058	0.126	118%	0.152	165%

The PH-transform with $r = 0.92$ yields a ddf of

$$S_Y(u) = 0.1^{0.92} \times \left(\frac{2{,}000}{2{,}000 + u}\right)^{1.2\times 0.92}.$$

For any excess layer $[a, a+h)$, the expected loss to the layer is

$$\mathrm{E}[X_{(a,a+h]}] = \int_a^{a+h} 0.1 \times \left(\frac{2{,}000}{2{,}000 + u}\right)^{1.2} du,$$

and the risk-adjusted premium by using a PH-transform ($r = 0.92$) is

$$H_r[X_{(a,a+h]}] = \int_a^{a+h} 0.1^{0.92} \times \left(\frac{2{,}000}{2{,}000 + u}\right)^{1.2\times 0.92} du.$$

Risk-adjusted premiums for various layers are shown in Table 3. In Table 3 we also list the prices by using a slightly different $r = 0.90$. Note that the developed prices are sensitive to the index r.

2.3. *Increased Limits Ratemaking*

In commercial liability insurance, a policy generally covers a loss (it may include allocated loss adjustment expense) up to a specified maximum dollar amount that will be paid on any individual loss. In the U.S., it is general practice to publish rates for some standard limit called the basic limit (historically \$25,000, but now \$100,000). Increased limit rates are calculated by applying increased limit factors (ILFs). Without risk load, the increased limit factor is the expected loss at the increased limit divided by the expected loss at the basic limit. The increased limit factor with risk load is the sum of the expected loss and the risk load at the increased limit divided by the sum of the expected loss and the risk load at the basic limit:

$$ILF(\omega) = \frac{\mathrm{E}[X;\omega] + RL_{(0,\omega]}}{\mathrm{E}[X;100{,}000] + RL_{(0,100{,}000]}}.$$

It is widely felt that ILFs should satisfy the following conditions (see Rosenberg [26], Meyers [16], and Robbin [25]). They implicitly assume that insureds who buy different limits are nevertheless subject to the same loss distributions.

1. The relative loading with respect to the expected loss is higher for higher limits.
2. ILFs should produce the same price under any arbitrary division of layers.
3. The ILFs should exhibit a pattern of declining marginal increases as the limit of coverage is raised. In other words, when $x < y$,

$$ILF(x+h) - ILF(x) \geq ILF(y+h) - ILF(y).$$

In the U.S., many companies use the ILFs published by the Insurance Services Office (ISO). Traditionally, only the severity distribution is used when producing ILFs. Until the mid-1980s, ISO used the variance of the loss distribution to calculate risk loads, a method proposed by Robert Miccolis [18]. From the mid-1980s to the early 1990s, ISO used the standard deviation of the loss distribution to calculate risk loads (e.g., Feldblum [4]).

Meyers [16] presents a competitive market equilibrium approach, which yields a variance-based risk load method; however, some authors have questioned the appropriateness of the variance-based risk load method for the calculation of ILFs (e.g., Robbin [25]).

The following is an illustrative example to show how the PH-transform method can be used in increased limits ratemaking.

EXAMPLE 5 Assume that the claim severity distribution has a Pareto distribution with ddf

$$S_X(u) = \left(\frac{\lambda}{\lambda + u}\right)^{\alpha},$$

with $\lambda = 5{,}000$ and $\alpha = 1.1$. This is the same distribution used by Meyers, although he also considered parameter uncertainty.

Assume that, based on the current market premium structure, the actuary feels that (for illustration only) an index $r = 0.9$ provides an appropriate provision for parameter uncertainty. When using a Pareto severity distribution, there is a simple analytical formula for the ILFs:

$$ILF(\omega) = \frac{1 - \left(\dfrac{\lambda}{\lambda + \omega}\right)^{r\alpha - 1}}{1 - \left(\dfrac{\lambda}{\lambda + 100{,}000}\right)^{r\alpha - 1}}.$$

One can then easily calculate the increased limit factors at any limit (see Table 4).

2.4. *Some Properties of the PH-Mean*

For a single risk X and for $0 \leq r \leq 1$, the PH-mean has the following properties (see Wang [30]):

- $\mathrm{E}[X] \leq H_r[X] \leq \max[X]$. When r declines from one to zero, $H_r[X]$ increases from the expected loss, $\mathrm{E}[X]$, to the maximum possible loss, $\max[X]$.

TABLE 4

INCREASED LIMIT FACTORS USING PH-TRANSFORM

Policy Limit ω	Expected Loss $E[X;\omega]$	ILF Without RL	Risk Load	ILF With RL
100,000	13,124	1.00	2,333	1.00
250,000	16,255	1.24	3,796	1.30
500,000	18,484	1.41	5,132	1.53
750,000	19,726	1.50	6,000	1.66
1,000,000	20,579	1.57	6,653	1.76
2,000,000	22,543	1.71	8,343	2.00

- Scale and translation invariant: $H_r[aX+b] = aH_r[X]+b$, for $a,b \geq 0$.

- Sub-additivity: $H_r[X+Y] \leq H_r[X]+H_r[Y]$.

- Layer additivity: when a single risk X is split into a number of layers

$$\{(x_0,x_1],(x_1,x_2],\ldots\},$$

the layer premiums are additive (the whole is the sum of the parts):

$$H_r[X] = H_r[X_{(x_0,x_1]}] + H_r[X_{(x_1,x_2]}] + \cdots.$$

Pricing often assumes that a certain degree of diversification will be reached through market efforts. In real life examples, risk-pooling is a common phenomenon. It is assumed that, in a competitive market, the benefit of risk-pooling is transferred back to the policyholders (in the form of premium reduction). In the PH-model, the layer-additivity and the scale-invariance have already taken into account the effect of risk-pooling. To illustrate, consider a single risk with a maximum possible loss of \$100 million. Suppose one insurer is asked to quote premium rates for each of the following as stand-alone coverages: sub-layers $(0,10], (10,20], \ldots, (90,100]$ and the whole risk $(0,100]$. The quoted premium for the entire risk $(0,100]$ may exceed the sum of individual premiums for each sub-layer. This is because the limit of $100 million may be a lot for a single insurer to carry without a substantial profit margin. However, the market mechanism would facilitate risk-sharing schemes among several insurers (say, ten insurers each take a sub-layer). Thus, when this risk pooling effect is transferred back to the policyholder, the premiums should be additive for different layers. Likewise, if one insurer is asked to quote premium rates for a 10% quota-share of this risk as opposed to the whole risk, the quoted premiums may exhibit non-linearity. However, the market risk-sharing scheme would force the premiums to be scale invariant—i.e., a 10% quota share demands 10% of the total premium.

Theoretically, in an efficient market (no transaction expenses in risk-sharing schemes) with complete information, the optimal cooperation among insurers is to form a market insurance portfolio (like the Dow Jones index), and each insurer takes a layer or quota-share of the market insurance portfolio.

In real life, however, the insurance market is *not* efficient. This is mainly because of incomplete information (ambiguity) and extra expenses associated with the risk-sharing transactions. There exist distinctly different local market climates in different geographic areas and in different lines of insurance. Catastrophe risk varies from region to region. In some geographic regions, due to high concentration and lack of information (ambiguity), existing risk-sharing schemes are not sufficient to diversify the risk to the extent one would wish. As a result the market would demand a higher risk load (a smaller value of r in the PH-model).

In summary, the index r may vary with respect to the local market climate which is characterized by *the levels of ambiguity*, *risk concentration*, and *competition*.

3. PRICING RISK PORTFOLIOS AND REINSURANCE TREATIES

When pricing a (re)insurance contract that covers a group of risks, the actuary often estimates claim frequency and claim severity separately, due to the type of information available. One straight-forward approach is to apply PH-transforms to the frequency and severity distributions separately, and then take the product of the loaded frequency and the loaded severity. An alternative is to first calculate the aggregate loss distribution from the estimated frequency and severity distributions, and then apply the PH-transform to the aggregate loss distribution. This section will discuss and compare both approaches.

3.1. Frequency/Severity Approach to Pricing Group Insurance

Let N denote the claim frequency with probability function $p_k = \Pr\{N = k\}$ and ddf $S_N(k) = p_{k+1} + p_{k+2} + \cdots$, $(k = 0, 1, 2, \ldots)$. The PH-mean for the frequency can be calculated as the sum

$$H_r[N] = S_N(0)^r + S_N(1)^r + S_N(2)^r + \cdots,$$

where convergence is required if N is unlimited (e.g., a Poisson frequency is unlimited).

Depending on the available information, the actuary may have different levels of confidence in the estimates for the frequency and severity distributions. According to the level of confidence in the estimated frequency and severity distributions, the actuary can choose an index r_1 for the frequency and an index r_2 for the severity. As a result, the actuary can calculate the risk-adjusted premium for the risk portfolio as

$$H_{r_1}[N] \times H_{r_2}[X].$$

EXAMPLE 6 Consider a group coverage of liability insurance. The actuary has estimated the following loss distributions: (i) the claim frequency has a Poisson distribution with $\lambda = 2.0$, and (ii) the claim severity is modeled by a lognormal distribution with a mean of \$50,000 and coefficient of variation of 3 (which was used by Finger [5] for a liability claim severity distribution). Here we also assume a coverage limit of one million dollars per claim. Suppose that the actuary has low confidence in the estimate of claim frequency, but higher confidence in the estimate of the claim severity distribution, and thus chooses $r_1 = 0.85$ for the claim frequency and $r_2 = 0.9$ for the claim severity. The premium can be calculated using numerical integrations:

$$H_{0.85}[N] = 2.227, \qquad \text{and} \qquad H_{0.9}[X] = 58{,}080.$$

Thus, the required total premium is

$$H_{0.85}[N] \times H_{0.9}[X] = 129{,}344.$$

Kunreuther et al. [13] discussed the ambiguities associated with the estimates for claim frequencies and severities. They mentioned that for some risks such as playground accidents, there are considerable data on the chances of occurrence but much uncertainty about the potential size of the loss due to arbitrary court awards. On the other hand, for some risks such as satellite losses or new product defects, the chance of a loss occurring is highly ambiguous due to limited past claim data. However, the magnitude of such a loss is reasonably predictable.

3.2. Frequency/Severity Approach to Pricing Per Risk Excess-of-Loss Reinsurance Treaties

A reinsurance excess-of-loss treaty normally covers a block of underlying policies where the attachment point and the policy limit apply on a per risk basis. For such reinsurance treaties, the claim frequency usually has a non-Bernoulli type distribution—that is, the number of claims may exceed one. For some low limit working layers where a substantial number of claims is expected, the major uncertainty might be in the claim frequency rather than in the severity.

In the market, reinsurance brokers often structure the coverage in a number of layers. It is important to have consistent pricing on all layers. Here we give an example.

EXAMPLE 7 Consider a reinsurance excess-of-loss treaty. The projected ceding company subject earned premium (SEP) for the treaty is \$10,000,000. The actuary is asked to price the following excess layers which are all on a per risk basis:

1. \$400K xs \$100K,
2. \$500K xs \$500K, and
3. The combined layer \$900K xs \$100K.

Suppose that, based on past loss data of the ceding company, after appropriate trending and development, the actuary has come up with the following best-estimates:

- The number of claims which cut into the first layer has a Poisson distribution with mean $\lambda = 6$.
- The size of losses greater than \$100K can be modeled by a single parameter Pareto with ddf

$$S_X(u) = \left(\frac{100}{u}\right)^{1.647}, \qquad u > 100.$$

Under this ground-up severity distribution, the loss to the first layer has a Pareto $(100, 1.647)$ distribution truncated at 400 with a ddf of

$$S_1(u) = \begin{cases} \left(\dfrac{100}{u+100}\right)^{1.647}, & 0 < u < 400; \\ 0, & 400 \leq u, \end{cases}$$

which has a mean severity of \$100,001.

The loss to the second layer has a Pareto (500, 1.647) distribution truncated at 500 with a ddf of

$$S_2(u) = \begin{cases} \left(\dfrac{500}{u+500}\right)^{1.647}, & 0 < u \leq 500; \\ 0, & 500 \leq u, \end{cases}$$

which has a mean severity of \$279,284.

In general, the frequency and severity distributions both change with the attachment point. To ensure consistency, it is important to work with the frequency and severity distributions for losses above the *minimum* attachment point. For convenience we refer to them as "ground-up" distributions, although they are not "real" ground-up distributions. In practice, reinsurers are usually supplied with data of large losses only, the "real" ground-up loss distribution below the attachment point is seldom known to the reinsurer.

By transforming the "ground up" frequency and severity distributions separately, we can load for the different frequency/severity risks accordingly.

For numerical illustration, we use the same PH-index $r = 0.95$ for both frequency and severity.

The PH-mean for a Poisson(6) distribution is 6.119, which represents a 1.98% frequency loading. First, we apply the PH-transform ($r = 0.95$) to the "ground-up" severity distribution and allocate the loaded costs to each layer (see Table 5). Under the Pareto (100, 1.647) "ground up" severity curve, the average severity in the layer 500 xs 500 is \$279,284, and the probability of cutting into the second layer given that a loss has cut into the first layer is 0.706. Therefore, the average loss to the second layer, *among all claims that have cut into the first layer*, can be calculated as the product $0.706 \times \$279,284 = \$19,717$.

TABLE 5

TRANSFORMING THE "GROUND UP" SEVERITY DISTRIBUTION

	Layer	Average Loss to the Layer Before Transform	Average Loss to the Layer After Transform $r = 0.95$	Relative Loading Ratio
(1)	400 xs 100	\$100,001	\$105,726	1.057
(2)	500 xs 500	\$ 19,717	\$ 23,117	1.172
(3)	900 xs 100	\$119,718	\$128,843	1.076
(1) + (2)	900 xs 100	\$119,718	\$128,843	1.076

TABLE 6

COMBINING LOADED "GROUND UP" FREQUENCY AND SEVERITY (PH-INDEX $r = 0.95$)

	Layer	Burning Cost (expected loss) As % of SEP	Loaded Rate As % of SEP $H_r[N] \times H_r[X]$	Relative Loading Ratio
(1)	400 xs 100	6.000%	6.469%	1.078
(2)	500 xs 500	1.183%	1.414%	1.196
(3)	900 xs 100	7.183%	7.883%	1.098
(1) + (2)	900 xs 100	7.183%	7.883%	1.098

Finally, we multiply the loaded "ground up" frequency and loaded severity in each layer to get the premium rate for each reinsurance layer (see Table 6). As a convention in reinsurance, the burning costs and premium rates are expressed as a percentage of the subject earned premium (SEP), \$10,000,000 in this example.

Note that, with this approach, we get premiums that are layer-additive. In other words, the total premium would not change regardless of how we divide the coverage into layers.

3.3. Aggregate Approach to Pricing Per-Risk Excess Treaty

As an alternative approach, the actuary can calculate/simulate the aggregate loss distribution from the best-estimate frequency and severity distributions, and subsequently apply the PH-transform to the aggregate loss distribution.

For given frequency N and severity X, let

$$Z = X_1 + X_2 + \cdots X_N$$

represent the aggregate loss amount for the risk portfolio. Various numerical and simulation techniques are available for calculating the aggregate loss distribution (e.g., Heckman and Meyers [7], and Panjer [19]).

In general, we get different results by transforming the frequency and severity distributions separately *versus* transforming the aggregate loss distribution. For the collective risk model, where claim severities are assumed mutually independent and independent of the frequency, we have the following inequality:

$$H_r[Z] \leq H_r[N] H_r[X], \qquad 0 < r < 1.$$

This is because, conditional on $N = n$, we always have

$$H_r[Z \mid N = n] = H_r[X_1 + \cdots + X_n] \leq n H_r[X].$$

In other words, the PH-transform of the aggregate loss distribution takes account of the fact that the variability regarding the aggregate loss is reduced in the pooling of N independent losses. However, one should carefully examine the validity of the independence assumption, especially with the presence of ambiguity (parameter uncertainties) in the best-estimate loss distributions. Parameter uncertainty can generate some correlation effect, although the claim processes may be independent provided that the true underlying distributions are known.

EXAMPLE 7 revisited: Now we re-consider the reinsurance treaty example using an aggregate approach. For ease of computation,

TABLE 7

APPLY PH-TRANSFORM TO THE AGGREGATE LOSS DISTRIBUTIONS OF EACH PER RISK EXCESS LAYER ($r = 0.9025$)

	Layer in 000's	Burning Cost As % of SEP	Indicated Rate As % of SEP	Relative Loading Ratio
(1)	400 xs 100	6.000%	6.384%	1.064
(2)	500 xs 500	1.183%	1.408%	1.190
(3)	900 xs 100	7.183%	7.742%	1.078
(1) + (2)	900 xs 100	7.183%	7.792%	1.085

here we assume independence among the individual claims in the calculation of the aggregate loss distribution for each layer. For a numerical comparison with the separate adjustment of frequency and severity, we apply a PH-transform with an index $r = r_1 \times r_2 = 0.95 \times 0.95 = 0.9025$ to the aggregate loss distribution of each layer. The indicated rate for each layer is given in Table 7.

We give some modeling details regarding this specific example. The claim frequency for the upper layer 500 xs 500 has a Poisson distribution with mean 0.424. This can be derived from the Poisson frequency for the lower layer 400 xs 100 and the probability of cutting into the second layer given that a loss has already cut into the first layer. Recall that the claim severity distribution for the layer 500 xs 500 has a Pareto (500, 1.647) distribution truncated at the policy limit 500. This can be verified using a conditional probability argument.

In this aggregate approach, we used a more severe PH-index $r = 0.9025$ as compared to $r = 0.95$. The aggregate approach produces a premium structure similar to that obtained by transforming frequency and severity separately (see Table 7 and Table 6). The use of a more severe index offsets the risk reduction as a result of pooling independent loss sizes.

Another important observation can be made from Table 7. With the aggregate approach, the premium rates are *not* additive for layers. The premium rate for the first layer (6.384%) plus that for the second layer (1.408%) is 7.792%, which is greater than the rate for the combined layer (7.742%). This lack of layer additivity may be a drawback of the aggregate approach in pricing per risk excess reinsurance treaties.

3.4. Aggregate Approach to Pricing Aggregate Contracts

Some reinsurance contracts are written in aggregate terms where the coverage triggers when the aggregate loss (or loss ratio) for the whole book exceeds some specified amount. Usually these contracts specify the attachment point and coverage limit in aggregate terms. In pricing such aggregate treaties, a natural approach would be to use the aggregate loss distribution, simply because the coverage trigger is the aggregate loss amount. In other words, the actuary needs to calculate/simulate a probability distribution for the aggregate loss $Z = X_1 + \cdots + X_N$. Based on the claim generating mechanism as well as the level of ambiguity, the actuary may assume some correlation between individual risks. The PH-transform of the aggregate loss distribution will automatically take into account the effect of correlation. The higher the correlation between individual risks, the greater the PH-mean for the aggregate loss distribution.

For some CAT events it might be plausible to consider the dependency between claim frequency and claim severity. For instance, the Richter scale value of an earthquake may affect both the frequency and severity simultaneously, and for hurricane losses, the wind velocity would affect both the frequency and severity simultaneously. Regardless of the dependency structure, computer simulation methods can always be used to model the aggregate losses based on a given geographic concentration. The PH-transform of the aggregate loss distribution can capture the correlation risk in the developed prices.

4. MIXTURE OF PH-TRANSFORMS

While a single index PH-transform has one parameter r to control the relative premium structure, one can obtain more flexible premium structures by using a mixture of PH-transforms:

$$p_1 H_{r_1} + p_2 H_{r_2} + \cdots + p_n H_{r_n},$$

$$\sum_{j=1}^{n} p_j = 1, \qquad 0 \le r_j \le 1 \quad (j = 1, \ldots, n).$$

The PH-index mixture can be interpreted as a collective decision-making process. Each member of the decision-making 'committee' chooses a value of r, and the index mixture represents different rs chosen by different members. It also has interpretations as (i) an index mixture chosen by a rating agency according to the indices for all insurance companies in the market; (ii) an index mixture which combines an individual company's index with the rating agency's index mixture.

A mixture of PH-transforms has the same properties as that for a single index PH-transform (see Section 2.4). For ratemaking purposes, a mixture of PH-transforms enjoys more flexibility than a single index PH-transform. Now we shall discuss two special mixtures of the form

$$(1-\alpha) H_{r_1}[X] + \alpha H_{r_2}[X], \qquad 0 \le \alpha \le 1, \quad r_1, r_2 \le 1.$$

4.1. Minimum Rate-on-Line

In most practical circumstances, very limited information is available for claims at extremely high layers. In such highly ambiguous circumstances, most reinsurers adopt a survival rule of minimum rate-on-line. The rate-on-line is the premium divided by the coverage limit, and most reinsurers establish a minimum they will accept for this ratio (see Venter [28]).

TABLE 8

LAYER PREMIUMS UNDER AN INDEX MIXTURE

Layer $X_{(a,b]}$ 1K=$1,000	Expected Loss	Risk-adjusted Premium	Percentage Loading
(0, 1K]	77.892	131.56	45.8%
(5K, 6K]	20.512	47.43	131%
(10K, 11K]	11.098	35.59	220%
(50K, 51K]	1.982	23.20	1,070%
(100K, 101K]	0.888	21.53	2,324%
(500K, 501K]	0.132	20.26	15,276%
(1,000K, 1,001K]	0.058	20.12	34,875%

By using a two-point mixture of PH-transforms with $r_1 \leq 1$ and $r_2 = 0$, the premium functional

$$(1-\alpha)H_{r_1}[X] + \alpha H_0[X] = (1-\alpha)H_{r_1}[X] + \alpha \max[X]$$

can yield a minimum rate-on-line at α.

EXAMPLE 8 Reconsider Example 4. The best-estimate loss distribution (ddf) is

$$S_X(u) = 0.1 \times \left(\frac{2{,}000}{2{,}000+u}\right)^{1.2}.$$

By choosing a two-point mixture with $r_1 = 0.92$, $r_2 = 0$, and $\alpha = 0.02$, we get an adjusted distribution:

$$S_Y(u) = 0.98 \times 0.1^{0.92} \times \left(\frac{2{,}000}{2{,}000+u}\right)^{1.2\times 0.92} + 0.02.$$

Note that S_Y being a proper loss distribution requires a finite upper layer limit.

As shown in Table 8, this two-point mixture guarantees a minimum rate-on-line at 0.02 (1 full payment out of 50 years). Note that the average index $\bar{r} = (1-\alpha)r_1 + \alpha r_2 = 0.9016 \approx 0.90$.

We can see that this method yields distinctly different premiums from those in Table 3 where the single indices $r = 0.92$ and $r = 0.90$ are used.

4.2. *The Right-Tail Deviation*

Consider a two-point mixture of PH-transforms with $r_1 = 1$ and $r_2 = \frac{1}{2}$:

$$(1-\alpha)H_1[X] + \alpha H_{1/2}[X] = \mathrm{E}[X] + \alpha(H_{1/2}[X] - \mathrm{E}[X]), \qquad 0 < \alpha < 1,$$

which is similar in form to the standard deviation method of $\mathrm{E}[X] + \alpha\sigma[X]$.

Now we introduce a new risk-measure analogous to the standard deviation.

Definition 2 *The right-tail deviation is defined as*

$$D[X] = H_{1/2}[X] - \mathrm{E}[X] = \int_0^\infty \sqrt{S_X(u)}\,du - \int_0^\infty S_X(u)\,du.$$

Analogous to the standard deviation, the right-tail deviation $D[X]$ satisfies the following properties:

- If $\Pr\{X = b\} = 1$, then $D[X] = 0$.
- Scale-invariant: $D[cX] = cD[X]$ for $c > 0$.
- Shift-invariant: $D[X + b] = D[X]$ for any constant b.
- Sub-additivity: $D[X + Y] \leq D[X] + D[Y]$.
- If X and Y are perfectly correlated, then $D[X + Y] = D[X] + D[Y]$.

It is shown in Appendix A that, for a small layer $(a, a+h]$, $D[X_{(a,a+h]}] \leq \sigma[X_{(a,a+h]}]$, and $D[X_{(a,a+h]}]$ converges to $\sigma[X_{(a,a+h]}]$

TABLE 9

THE RIGHT-TAIL DEVIATION VERSUS THE STANDARD DEVIATION

Layer 1K = $1,000	Expected loss	Std-deviation of the loss	Right-tail deviation	Percentage difference
L	$\mathrm{E}[L]$	$\sigma[L]$	$D[L]$	$\frac{\sigma[L]}{D[L]} - 1$
(0, 1K]	77.89	256.0	200.5	27.7%
(1K, 2K]	51.56	214.3	175.2	22.3%
(10K, 11K]	11.10	103.9	94.24	10.3%
(100K, 101K]	.8879	29.76	28.91	2.93%
(1,000K, 1,001K]	.05754	7.584	7.528	0.75%
(10,000K, 10,001K]	.003640	1.908	1.904	0.19%

at upper layers (i.e., the relative error goes to zero when a becomes large). As a result, the right-tail deviation $D[X]$ is finite if and only if the standard deviation $\sigma[X]$ is finite.

EXAMPLE 9 Re-consider the loss distribution in Example 4 with a ddf of

$$S_X(u) = 0.1 \times \left(\frac{2{,}000}{2{,}000 + u}\right)^{1.2}.$$

For different layers with fixed limit at 1000, Table 9 compares the standard deviation with the right-tail deviation.

Having stated a number of similarities, here we point out two crucial differences between the right-tail deviation $D[X]$ and the standard deviation $\sigma[X]$ (see Wang [31]):

- $D[X]$ is layer-additive, but $\sigma[X]$ is *not* additive.
- $D[X]$ preserves some natural ordering of risks such as first stochastic dominance,[2] but $\sigma[X]$ does not.

[2]Risk X is smaller than risk Y in first stochastic dominance if $S_X(u) \le S_Y(u)$ for all $u \ge 0$; or equivalently, Y has the same distribution as $X + Z$ where Z is another non-negative random variable.

TABLE 10

LAYER ADDITIVITY: A COMPARISON

Layer	Expected Losses	Standard Deviation	Right-Tail Deviation
(0, 10K]	301	1378	1355
(10K, 20K]	80	834	809
(0, 20K]	381	2035	2164
Result	Additive	Sub-Additive	Additive

These two crucial differences give the right-tail deviation an advantage over the standard deviation in pricing insurance risks.

Although for a small layer $(a, a+h]$ we have $D[X_{(a,a+h]}] \leq \sigma[X_{(a,a+h]}]$, for the entire risk X the right-tail deviation often exceeds the standard deviation, since the right-tail deviation is layer-additive while the standard deviation is not. For example, consider two sub-layers $(0, 10\text{K}]$ and $(10\text{K}, 20\text{K}]$, and a combined layer $(0, 20\text{K}]$. The right-tail deviation exceeds the standard deviation for the combined layer (0, 20K], although the reverse relation holds for each sub-layer (see Table 10).

Remark For a layer $(a, a+h]$, the loss-to-limit ratio is defined as the ratio of incurred loss to the limit of the layer. When the layers are refined (h becomes small), the loss-to-limit ratio approaches the ddf at that layer, which is also the frequency of hitting the layer. This can be seen from the relation

$$\lim_{h \to 0} \frac{\int_a^{a+h} S_X(u)\,du}{h} = S_X(a).$$

If a ground-up risk is divided into small adjacent layers, the empirical loss-to-limit ratios at various layers yield an approximation to the underlying ddf. As a pragmatic method for computing risk loads, it has been a longstanding practice of some reinsurers to adjust the empirical loss-to-limit ratio by adding a multiple of the square root of the empirical loss-to-limit ratio. As the layers are refined (h becomes small), this pragmatic method approaches the following:

$$\mathrm{E}[X_{(a,a+h]}] + \alpha(D[X_{(a,a+h]}] + \mathrm{E}[X_{(a,a+h]}]).$$

5. ECONOMIC THEORIES OF RISK LOAD

In this section we review some economic theories and show how the PH-transform fits in.

5.1. Expected Utility Theory

Traditionally, expected utility (EU) theory has played a dominant role in modeling decisions under risk and uncertainty. To a large extent, the popularity of EU was attributed to the axioms of von Neumann and Morgenstern [29].

Let V represent a random economic prospect and let $S_V(u) = \Pr\{V > u\}$ (i.e., the probability that the random economic prospect V exceeds value u). Let the symbols $\succ$ and $\sim$ represent strict preference and indifference, respectively. Von Neumann and Morgenstern proposed five axioms of decision under risk:

- EU.1 If V_1 and V_2 have the same probability distribution, then $V_1 \sim V_2$.
- EU.2 Weak order: $\succeq$ is reflective, transitive, and connected.
- EU.3 $\succeq$ is continuous in the topology of weak convergence.
- EU.4 If $V_1 \geq V_2$ with probability one, then $V_1 \succeq V_2$.
- EU.5 If $S_{V_1} \succeq S_{V_2}$, and for any $p \in [0,1]$, the probabilistic mixture satisfies

$$(1-p)S_{V_1} + pS_{V_3} \succeq (1-p)S_{V_2} + pS_{V_3}.$$

Von Neumann and Morgenstern showed that any decision-making which is consistent with these five axioms can be modeled by a utility function u such that '$V_1 \succ V_2$ if and only if $\mathrm{E}[u(V_1)] \geq \mathrm{E}[u(V_2)]$.'

When EU is applied to produce an insurance premium for a risk X, the minimum premium P that an insurance company will accept for full insurance satisfies the EU-equation

$$u(w) = \mathrm{E}[u(w + P - X)]$$

in which u and w refer to the insurer's utility function and wealth (see Bowers et al. [2]). The premium P from the EU-equation does not satisfy layer-additivity. Thus, the PH-transform does not fit in the expected utility framework.

5.2. The Dual Theory of Yaari

Modern economic theory questions the assumption that a firm can have a utility function, even when it accepts that individuals do. Yaari [32] proposed an alternative theory of decision under risk and uncertainty.

While the first four EU axioms are apparently reasonable, many people challenged the fifth axiom in the expected utility theory. While keeping the first four EU axioms unchanged, Yaari proposed an alternative to the fifth EU axiom:

DU.5* If V_1, V_2, and V_3 are co-monotone and $V_1 \succeq V_2$, for any $p \in [0,1]$, the outcome mixture satisfies

$$(1-p)V_1 + pV_3 \succeq (1-p)V_2 + pV_3.$$

Two risks X and Y are *co-monotone* if there exists a random variable Z and non-decreasing real functions u and v such that $X = u(Z)$ and $Y = v(Z)$ with probability one. Co-monotonicity is a generalization of the concept of perfect correlation to random variables without linear relationships. Note that perfectly correlated risks are co-monotone, but the converse does not hold. Consider two layers $(a, a+h]$ and $(b, b+h]$ for a continuous variate X. The layer payments $X_{(a,a+h]}$ and $X_{(b,b+h]}$ are co-monotone since both are non-decreasing functions of the original risk X. They are bets on the same event and neither of them is a *hedge* against the other. On the other hand, for $a \neq b$, $X_{(a,a+h]}$ and $X_{(b,b+h]}$ are *not* perfectly correlated since neither can be written as a linear function of the other.

Under axioms EU.1–4 & EU.5*, Yaari showed that there exists a distortion function $g : [0,1] \to [0,1]$ such that a certainty equivalent to a random economic prospect V on interval $[0,1]$ is

$$\int_0^1 g[S_V(y)]\,dy.$$

In other words, the certainty equivalent to a random economic prospect, $0 \leq V \leq 1$, is just the expected value under the distorted probability distribution, $g[S_V(y)]$, $0 \leq y \leq 1$.

Regarding the concept of risk-aversion, Yaari made the following observations:

> At the level of fundamental principles, risk-aversion and diminishing marginal utility of wealth, which are synonymous under expected utility theory, are horses of different colors. The former expresses an attitude towards risk (increased uncertainty hurts) while the latter expresses an attitude towards wealth (the loss of a sheep hurts more when the agent is poor than when the agent is rich). [32, p. 95]

The PH-transform fits in Yaari's economic theory with $g(x) = x^r$.

5.3. *Schmeidler's Ambiguity-Aversion*

As early as 1921, John Keynes identified a distinction between the *implication* of evidence (the implied likelihood) and the *weight* of evidence (confidence in the implied likelihood). Frank Knight [10] also drew a distinction between *risk* (with known probabilities) and *uncertainty* (ambiguity about the probabilities). A famous example on ambiguity-aversion is Ellsberg's [3] paradox which can be briefly described as follows: There are two urns each containing 100 balls. One is a non-ambiguous urn which has 50 red and 50 black balls; the other is an ambiguous urn which also contains red and black balls but with unknown proportions. When subjects are offered \$100 for betting on a red draw, most subjects choose the non-ambiguous urn (and the same for the black draw). Such a pattern of preference *cannot* be explained by EU (Quiggin, [24, p. 42]).

Ellsberg's work has spurred much interest in dealing with ambiguity factors in risk analysis. Schmeidler [27] brought to economists *non-additive probabilities* in his axiomization of preferences under uncertainty. For instance, in Ellsberg's experiment, the non-ambiguous urn with 50 red and 50 black balls is preferred to the ambiguous urn with unknown proportions of red or black balls. This phenomenon can be explained if we assume that one assigns a subjective probability $\frac{3}{7}$ to the chance of getting a red draw (or black draw). Since $\frac{3}{7} + \frac{3}{7} = \frac{6}{7}$ which is less than one, the difference $1 - \frac{6}{7} = \frac{1}{7}$ may represent the magnitude of ambiguity aversion.

In his axiomization of acts or risk preferences, Schmeidler obtained essentially the same mathematical formulation (axioms and theorems) as that of Yaari. A certainty equivalent to a random economic prospect V $(0 \leq V \leq 1)$ can be evaluated as

$$H[V] = \int_0^1 g[S_V(u)]\, du,$$

where $g : [0,1] \mapsto [0,1]$ is a distortion (increasing, non-negative) function, and $g[S_X(u)]$ represents the subjective probabilities.

The major difference between the Schmeidler model and the Yaari model lies in the interpretation (Quiggin, [24]). Yaari assumes that the objective distributions (e.g., S_X) are known and one applies a distortion (i.e., g) to the objective distribution. Schmeidler argues that it is illogical to assume an objective distribution; instead, he interprets $g \circ S_X$ as non-additive subjective probabilities which can be inferred from acts.

However, economic interpretations are important. For instance, if the underlying distribution is assumed to be known, then the process risks can be diversified away in a risk portfolio. Ambiguity is uncertainty regarding the best-estimate probability distribution, and thus may not be diversifiable in a risk portfolio.

The PH-transform fits in Schmeidler's economic theory with an interpretation of aversion to ambiguity (parameter risk).

5.4. No-Arbitrage Theory of Pricing

No-arbitrage is a fundamental principle in financial economic theory, which requires linearity of prices. The theories of Yaari and Schmeidler can be viewed as a more relaxed (or more general) version of the no-arbitrage theory, i.e., they only require no arbitrage (linearity) on co-monotone risks (e.g., different layers of the same risk). Using a market argument, Venter [28] discussed the no-arbitrage implications of reinsurance pricing. He observed that in order to ensure additivity when layering a risk, it is necessary to adjust the loss distribution so that layer premiums are expected losses under the adjusted loss distribution. Venter's observation is in agreement with the theories of Yaari and Schmeidler. In fact, the PH-transform is a specific transform which conforms to Venter's adjusted distribution method.

6. SUMMARY

In this paper we have introduced the basic methodologies of the PH-transform method and have shown by example how it can be used in insurance ratemaking. We did not discuss how to decide the overall level of contingency margin, which depends greatly on market conditions. An important avenue for future research is to link the overall level of risk load with the required surplus for supporting the written contract. Some pioneering work in this direction can be found in Kreps [11], [12] and Philbrick [22].

The use of adjusted/conservative life tables has long been practiced by life actuaries (see Venter, [28]). To casualty actuaries, the PH-transform method contributes a theoretically sound and practically plausible way to adjust the loss distributions.

REFERENCES

[1] Bault, Todd, Discussion of Feldblum: "Risk Loads for Insurers," *PCAS* LXXXII, pp. 78–96.

[2] Bowers, Newton L., Hans U. Gerber, James C. Hickman, Donald A. Jones, and Cecil J. Nesbitt, *Actuarial Mathematics*, Itasca, IL: Society of Actuaries, 1986.

[3] Ellsberg, Daniel, "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75, 1961, pp. 643–669.

[4] Feldblum, Sholom, "Risk Loads for Insurers," *PCAS* LXXVII, 1990, pp. 160–195.

[5] Finger, Robert J., "Estimating Pure Premiums by Layer—An Approach," *PCAS* LXIII, 1976, pp. 34–52.

[6] Hogarth, Robin M. and Howard Kunreuther, "Price Insurance and Warranties: Ambiguity and Correlated Risks," *The Geneva Papers on Risk and Insurance Theory*, 17, 1, 1992, pp. 35–60.

[7] Heckman, Philip and Glenn Meyers, "The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions," *PCAS* LXX, 1983, pp. 22–61.

[8] Hogg, Robert and Stuart Klugman, *Loss Distributions*, John Wiley, New York, 1984.

[9] Keynes, John M., *A Treatise on Probability*, London: Macmillian, 1921.

[10] Knight, Frank H., *Risk, Uncertainty and Profit*, Boston: Houghton Mifflin, 1921.

[11] Kreps, Rodney, "Reinsurer Risk Loads from Marginal Surplus Requirements," *PCAS* LXXVII, 1990, pp. 196–203.

[12] Kreps, Rodney, "Investment-Equivalent Reinsurance Pricing," *PCAS* LXXXV, 1998, pp. 101–131.

[13] Kunreuther, Howard, Robin Hogarth, and Jacqueline Meszaros, "Insurer Ambiguity and Market Failure," *Journal of Risk and Uncertainty*, 7, 1993, pp. 71–87.

[14] Lawless, Jerry F., *Statistical Models and Methods For Lifetime Data*, John Wiley & Sons, New York, 1982.

[15] McClenahan, Charles L., "Ratemaking," in *Foundations of Casualty Actuarial Science*, Chapter 2, Casualty Actuarial Society, New York, 1990.

[16] Meyers, Glenn G., "The Competitive Market Equilibrium Risk Load Formula for Increased Limits Ratemaking," *PCAS* LXXVIII, 1991, pp. 163–200.

[17] Meyers, Glenn G., "Managing the Catastrophe Risk," *Incorporating Risk Factors in Dynamic Financial Analysis*, Casualty Actuarial Society Discussion Paper Program, May 1995, pp. 111–150.

[18] Miccolis, Robert S., "On the Theory of Increased Limits and Excess-of-Loss Pricing," *PCAS* LXIV, 1977, pp. 27–59.

[19] Panjer, Harry H., "Recursive Evaluation of a Family of Compound Distributions," *ASTIN Bulletin*, 12, 1981, pp. 22–26.

[20] Patrick, S. Gary, "Reinsurance," in *Foundations of Casualty Actuarial Science*, Chapter 6, Casualty Actuarial Society, New York, 1990.

[21] Philbrick, Stephen, Discussion of Feldblum: "Risk Loads for Insurers," *PCAS* LXXVIII, 1991, pp. 56–63.

[22] Philbrick, Stephen W., "Accounting For Risk Margins," Casualty Actuarial Society *Forum*, 1, Spring 1994, pp. 4–87.

[23] Quiggin, John, "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 4, 1982, pp. 323–343.

[24] Quiggin, John, *Generalized Expected Utility Theory: The Rank-Dependent Model*, Kluwer Academic Publishers, Boston, 1993.

[25] Robbin, Ira, Discussion of Meyers: "The Competitive Market Equilibrium Risk Load Formula for Increased Limits Ratemaking," *PCAS* LXXIX, 1992, pp. 367–384.

[26] Rosenberg, Sheldon, Discussion of Miccolis: "On the Theory of Increased Limits and Excess-of-Loss Pricing," *PCAS* LXIV, 1977, pp. 60–73.

[27] Schmeidler, David, "Subjective Probability and Expected Utility Without Additivity," *Econometrica*, 57, 1989, pp. 571–587.

[28] Venter, Gary G., "Premium Calculation Implications of Reinsurance Without Arbitrage," *ASTIN Bulletin*, 21, 1991, pp. 223–230.

[29] von Neumann, John and Oskar Morgenstern, *Theory of Games and Economic Behavior*, (2nd edition), Princeton, NJ: Princeton University Press, 1947.

[30] Wang, Shaun, "Insurance Pricing and Increased Limits Ratemaking by Proportional Hazards Transforms," *Insurance Mathematics and Economics*, 17, 1995, pp. 43–54.

[31] Wang, Shaun, "Premium Calculation by Transforming the Layer Premium Density," *ASTIN Bulletin*, 26, 1, 1996, pp. 71–92.

[32] Yaari, Menahem E., "The Dual Theory of Choice Under Risk," *Econometrica*, 55, 1987, pp. 95–115.

APPENDIX A: PROOFS OF SOME STATED RESULTS

THEOREM 1 *For any non-negative random variable X (discrete, continuous, or mixed), we have*

$$\mathrm{E}[X] = \int_0^\infty S_X(u)\,du.$$

Proof For $x \geq 0$ it is true that

$$x = \int_0^\infty I(x > u)\,du,$$

where I is the indicator function (assuming values of 0 and 1). For a non-negative random variable it holds that

$$X = \int_0^\infty I(X > u)\,du.$$

By taking expectation on both sides of the equation one gets

$$\mathrm{E}[X] = \int_0^\infty \mathrm{E}[I(X > u)]\,du = \int_0^\infty S_X(u)\,du. \qquad \square$$

THEOREM 2 *For a small layer $[a, a+h)$ with h being a small positive number, we have*

- $D[X_{(a,a+h]}] \leq \sigma[X_{(a,a+h]}]$, *and*
- $\lim_{a\to\infty} D[X_{(a,a+h]}]/\sigma[X_{(a,a+h]}] = 1$.

Proof Let $p = S_X(u)$ be the probability of hitting the layer $[a, a+h)$. Note that $p \to 0$ as $u \to \infty$. The payment by the small layer $[a, a+h)$ has approximately a Bernoulli type distribution:

$$\Pr\{X_{(a,a+h]} = 0\} = 1 - p, \qquad \Pr\{X_{(a,a+h]} = h\} = p.$$

Thus, $\sigma(X_{(a,a+h]}) = \sqrt{p - p^2}h$ and $D[X_{(a,a+h]}] = (\sqrt{p} - p)h$. The two results come from the fact that $\sqrt{p} - p \leq \sqrt{p - p^2}$ for $0 \leq p \leq 1$ and

$$\lim_{p\to 0} \frac{\sqrt{p} - p}{\sqrt{p - p^2}} = 1. \qquad \square$$

APPENDIX B: AMBIGUITY AND PARAMETER RISK

Most insurance risks are characterized by the uncertainty about the estimate of the tail probabilities. This is often due to data sparsity for rare events (small tail probabilities), which in turn causes the estimates for tail probabilities to be unreliable.

To illustrate, assume that we have a finite sample of n observations from a class of identical insurance policies. The empirical estimate for the loss distribution is

$$\hat{S}(u) = \frac{\text{\# of observations} > u}{n}, \qquad u \geq 0.$$

Let $S(u)$ represent the *true* underlying loss distribution, which is generally unknown and different from the empirical estimation $\hat{S}(u)$. From statistical estimation theory (e.g., Lawless [14, p. 402], Hogg and Klugman [8]), for some specified value of u, we can treat the quantity

$$\frac{\hat{S}(u) - S(u)}{\sigma[\hat{S}(u)]},$$

as having a standard normal distribution for large values of n, where

$$\sigma[\hat{S}(u)] \approx \frac{\sqrt{\hat{S}(u)[1-\hat{S}(u)]}}{\sqrt{n}}.$$

The η% *upper confidence limit (UCL)* for the true underlying distribution $S(u)$ can be approximated by

$$UCL(u) = \hat{S}(u) + \frac{q_\eta}{\sqrt{n}}\sqrt{\hat{S}(u)[1-\hat{S}(u)]},$$

where q_η is a quantile of the standard normal distribution: $\Pr\{N(0,1) \leq q_\eta\} = \eta$. Keeping n fixed and letting $t \to \infty$, the ratio of the UCL to the best-estimate $\hat{S}(u)$ is

$$\frac{UCL(u)}{\hat{S}(u)} = 1 + \frac{q_\eta}{\sqrt{n}}\sqrt{\frac{1-\hat{S}(u)}{\hat{S}(u)}} \to \infty,$$

which grows without bound as u increases.

As a means of dealing with ambiguity regarding the best-estimate, the PH-transform

$$\hat{S}_Y(u) = [\hat{S}_X(u)]^r, \qquad r \leq 1,$$

can be viewed as an upper confidence limit (UCL) for the best-estimate $\hat{S}_X(u)$. It automatically gives higher relative safety margins for the tail probabilities, and the ratio

$$\frac{[\hat{S}_X(u)]^r}{\hat{S}_X(u)} = [\hat{S}_X(u)]^{r-1} \to \infty, \qquad \text{as} \quad u \to \infty,$$

increases without bound to infinity.