# Discussion of Dr. Shaun Wang's

# Implementation of PH-Transforms in Ratemaking

Discussion by Gary G. Venter

# Discussion of PH-Transforms

Dr. Wang has provided a good case fur the use of the mean of the PH-transform of a loss distribution as the risk-loaded premium. I would like to comment on several issues: 1) the need for consistency of the adjustment among contracts; 2) alternative transforms; 3) calibration; 4) the need for arbitrage free methods; 5) links to other premium loading methods; 6) minimum rates on line; 7) connecting to Yaari and Schmeidler.

## 1) Consistency among Contracts

In Venter (1991) I argued that the following aspects of a risk load formula are all equivalent:

1. It is arbitrage free.
2. It is strictly additive for both independent and correlated risks.
3. It can be calculated as the expected value of an adjusted probability distribution.

Note however that this does not guarantee a positive load over expected values. That depends on the adjusted distribution chosen and the cover being priced.

The PH-transform would be one choice for generating such an adjusted distribution. When considering various contracts and portfolios, however, the use of adjusted distributions produces additive and arbitrage-free pricing only if a single adjustment of probability is made for each event and this event probability is kept fixed when looking at the various contracts. That is, to avoid arbitrage, once a transformed probability has been selected for an event, that probability has to be used for the entire calculation.

This is violated, for example, in Wang's Table 7. He applies PH-transformed probabilities to the aggregate losses in the occurrence layers 400 xs 100 and 500 xs 500, and calculates the mean loss to each layer under those probabilities (6.384% and 1.408% respectively). Under those probabilities, the loss to the layer 900 xs 100 then has to be the sum, or 7.792%. This is regardless of independence – the means are always additive. But as Wang shows, the PH-transform of the 900 xs 100 layer using the same parameter gives a different mean, and so must imply different probabilities for the layer losses.

This reasoning may lead to inconsistencies any time the PH-transform is applied to aggregate losses, especially if any hypothetical contract has to be priced. This would suggest applying transformations separately to frequency and severity. In any case, for use in reinsurance pricing, it would be useful to have separate transforms of the frequency and severity distributions, so that the transformed distributions can be used in excess, proportional, and aggregate treaty pricing.

For this reason, and for ease of calculation, the PH-transform is probably best applied to just the severity distribution. Frequency can be adjusted more simply, perhaps by changing the parameters of the frequency distribution. The additivity that results from adjusting severity with a single transformation applied to the entire severity distribution is illustrated in Wang's Table 5.

The need for consistent adjustment of probabilities in order to produce additivity can also be illustrated using an example from Delbaen and Haezendonck (1989). One of the transforms they consider is the adjusted distribution $f^*(x) = f(x)[1+h(x-E(X))]$. This gives the adjusted mean $E^*(X) = E(X) + hVar(X)$. Thus they seem to have shown that a variance load is a form of adjusted probability. However, for this to be a true variance load, inconsistent probability adjustments are sometimes required. For instance, if Y is adjusted by the same method, the price for Y would be the adjusted mean $E^*(Y) = E(Y) + hVar(Y)$. These probability adjustments would then determine the adjusted mean for X + Y to be $E^*(X+Y) = E(X+Y) + hVar(X) + hVar(Y)$, as means are additive even for correlated variates. If X and Y are in fact correlated, this is not the same as the variance load $E(X+Y) + hVar(X+Y)$. This load could be achieved as an adjusted mean for X+Y, but different probabilities for X and Y would be needed. Thus for an adjusted mean to always give the variance load, the transformed probabilities must change during the calculation. If the variable X is kept fixed when transforming other variables, this transformation becomes a covariance load which produces the variance load only for X itself.

Even variables that depend only on X do not always get a variance loading under this transformation. For instance, a 2% quota share of X would have mean under $f^*$ of $E^*(0.02X) = 0.02[E(X) + hVar(X)]$, but under a variance load it would have a price of $0.02E(X) + h(0.02)^2Var(X)$. This is only $1/50^{th}$ as much loading, so with a variance load a risk could be 100% ceded to 50 reinsurers and 98% of the profit retained by the cedant. With the adjusted probability method the whole profit would go to the reinsurers.

## 2) An Alternative Transform

Buhlmann (1980) suggested using the Esscher transform to calculate risk load. This transformation is $f^*(x) = e^{hx}f(x)/E(e^{hX})$. More recently Gerber and Shiu (1993) suggest using the Esscher transform on ln X. They show that if you do that for lognormally distributed security prices, you get another lognormal with just an adjustment to the $\mu$ parameter. The transformed $\mu$ parameter is $\mu+h\sigma^2$. If you then calibrate h to produce the current security price as the discounted transformed mean of the original security, they show that you recover the Black-Scholes option pricing formula as the adjusted expected present value of the option. This makes the Esscher transform on logs interesting, as in this case it ties in with known financial theory. It can sometimes be calculated more simply as the equivalent power transform of the original distribution $f^*(x) = x^r f(x)/E(X^r)$.

However, for Poisson variables, they show that the Esscher transform of the original variable (not the log) agrees with option prices for jump processes. Here the transform is also Poisson. A reasonable approach for compound frequency / long-tailed severity processes may be to apply the Esscher transform to frequency and to the log of severity.

## 3) Calibration

Both the PH and Esscher transforms have a free constant that determines the level of the risk load. The value of this constant will depend on market conditions. However, if the price for basic primary coverage is known, it can be used to determine the value of the constant, which then can be used to price different reinsurance or excess layers for the same risk. This is basically the approach Black and Scholes use to price derivative contracts based on the pricing of the underlying asset.

As an example, suppose a portfolio of commercial property risks has a disappearing deductible of 1 (in appropriate units), and severity distribution $F(x) = 1 - x^{-a}$, i.e., the simple Pareto, for x>1. The distribution for ln X is the exponential distribution $G(x) = 1 - e^{-ax}$. For this, the moment generating function is $M(h) = E(e^{hx}) = a/(a-h)$. Then the Esscher transform on G is $g^*(x) = e^{hx}g(x)(a-h)/a$. Since $g(x) = ae^{-ax}$, this shows that $g^*(x) =$

$(a-h)e^{-(a-h)x}$, which is another exponential distribution. Thus $F^*(x) = 1 - x^{-(a-h)}$ is the transformed severity distribution. But this distribution can also be reached as a PH-transform of the original Pareto. In this case, then, the log-Esscher and PH-transform can be calibrated to give the same distribution.

To determine the constant h, consider what the loading is for the portfolio. The mean for the simple Pareto is a/(a-1). Suppose a=2, and a loading of 10% of premium is built into primary pricing. Assume no loading is made for frequency. Then E(X) = 2, and h is needed so that (a-h)/(a-h-1) = 2.2, or (2-h)/(1-h) = 2.2. This gives h = 1/6. Thus the Pareto with a=1.833 gives the primary price as its mean, and so can be used for consistent pricing of reinsurance layers. The corresponding r for the PH-transform would be r=11/12=0.9166, as $(x^{-2})^{0.9166} = x^{-1.833}$. Note that the transformed distribution is zero below x=1, as is the original. This is an important detail omitted from the above discussion: in order to avoid arbitrage the transformed distribution must give zero probability to the same events as does the original distribution.

A problem with this calculation is that a loading should be made for frequency, as the frequency risk will not be carried for free. Reinsurance contracts for aggregate losses or for aggregate coverage on small per-occurrence layers would carry substantial frequency risk, which would have to be priced. The above calculation of the severity r or h can be done for any assumed frequency load. Rather than no load, another assumption might be that the frequency and severity loads are the same, or are in some pre-selected proportion. Thus there is still some judgment involved in trying to find the adjusted probabilities that support the underlying pricing.

The PH and log-Esscher transforms are not always the same. For instance, for the inverse Weibull distribution $F(x) = \exp(-(\theta/x)^\tau)$, the PH-transform is $1-[1 - \exp(-(\theta/x)^\tau)\}^r$, whereas the log-Esscher transform is inverse transformed gamma. These will tend to be similar in typical cases, however, depending on the calibration. This suggests that the choice between PH and log-Esscher transforms for severity may be largely a matter of ease of calculation. For example, the log-Esscher transform of the lognormal is lognormal, whereas the PH-transform is more complicated.

## 4) Should Actuaries Always Use Arbitrage Free Methods?

Insurers may want to try to build arbitrage possibilities into prices, e.g., by using variance loadings. If these loadings succeed in the market, this might give the insurers arbitrage opportunities. Exploiting arbitrage opportunities is usually regarded as producing an improvement in market functioning, as it tends to compete away those opportunities. However, this policy would need to be monitored carefully, as the products with arbitrage might eventually lose out to other competitors, resulting in a loss of market share.

Some apparent arbitrage in the reinsurance market may not actually be such. For instance, a strongly capitalized insurer may cede a portion of its risk to a group of weaker reinsurers for less than it received in premium. But it is bearing the insolvency risk of those reinsurers, while its customers are not.

## 5) Links to Other Risk Load Methods

I was interested in Dr. Wang's comment that while the PH-transform aims to give the market price, different insurers may want different prices, depending for instance on their current risk portfolios. While he does not discuss how an insurer's desired price would be calculated, a logical approach might be to try to quantify the degree of risk assumed, and look to market norms or carrier goals to determine what the return should be for carrying that risk. This in general terms is how some other risk load methodologies proceed, e.g., the papers of Kreps and Meyers that Dr. Wang's paper cites. Thus those approaches could be considered to be aiming at an insurer's desired risk load, while the transformed distributions are looking for the market price.

## 6) Minimum Rates on Line

Wang shows that a mixture of PH-transforms can produce minimum rates on line as market prices for low-expected-value risks. For individual reinsurers minimum rates on line may make sense due to capacity constraints. They are problematic from a market theory viewpoint, however, for if a group of risks with very low expected losses were each written at a market minimum rate on line, they could conceivably be packaged and ceded as a group at that same rate on line, generating an arbitrage profit. It could be that market minimums exist as barriers but not as actual prices, as they may serve only to stop purchases when the expected loss is enough less than the rate. If so, then the above problem could not arise, as two or more minimum rate risks would cost more than the minimum when combined as a group.

Even without the mixture, a single PH or Esscher transform can produce layer rates that decline very slowly as the retention increases, and the risk load as a percentage of expected losses can increase without limit. Although not giving a true minimum rate on line, this could approximate market behavior fairly reasonably.

## 7) Connecting to Yaari and Schmeidler

Both of these approaches advocate pricing by the mean of a distortion of the ddf, denoted by Wang as $g[S_V(u)]$. This makes it look like the distorted probability should act directly on the ddf. However, any distortion of probability can be re-expressed to state its effect on the ddf. For instance, consider a scale transform of a Pareto with original ddf of $(1+u/b)^{-a}$ and transformed ddf of $(1+u/c)^{-a}$. If $g(x) = [1+b(x^{-1/a} – 1)/c]^{-a}$, then it will produce the transformed ddf from the original ddf.

Perhaps surprisingly, an increased scale parameter does not always produce a positive loading. An example where it does not is attributed to Thomas Mack in Albrecht (1992). It turns out that if you have a disappearing deductible, and want a separate cover to buy that back to full coverage, the buyback of the deductible can be cheaper for a higher scale parameter. For instance, suppose severity is Pareto with $S(x) = (1+x/b)^{-a}$, and you want to buy a cover that pays the full loss X up to X=c, and nothing above c. The expected loss for this cover is the expected loss limited to c, $E(X|c)$, less $cS(c)$. From $E(X|c) = [1 – (1+c/b)^{1-a}]b/(a – 1)$ and taking a=2, this simplifies to $bc^2/(b+c)^2$. This can be seen to be a decreasing function of b for $c < b$, so in that case an increased scale parameter would give an adjusted mean less than the actual mean. The scale transform is pushing probability up to higher loss levels, so there is less below the deductible.

Unfortunately, this can happen with the PH and log-Esscher transforms also. The PH-transform of a Pareto lowers the a parameter. For b=1 the deductible expected value is $[1 - (1+da)(1+d)^{-a}]/(a - 1)$, which is an increasing function of a in some ranges. Thus lowering a will lower the expected value, making the adjusted mean less than the true mean. The log-Esscher transform in h of the Pareto is the generalized Pareto distribution with parameters a – h, b, h+1. The deductible mean under this is also sometimes less than the non-transformed mean.

Even Delbaen and Haezendonck's loading of a portion of the covariance of the deductible with full coverage would have the same problem, as when c is low enough, the deductible is negatively correlated with total losses due to it being zero for larger loss amounts. For all of these approaches the full coverage contract would have less risk load than coverage excess of a small disappearing deductible.

This may be appropriate when the losses under the deductible are negatively correlated with total losses. When the variance of total losses is less than that of the losses excess of the disappearing deductible, even a traditional standard deviation or variance loading will price a buyback of the deductible for an excess policyholder at less than the expected losses of the additional coverage. For instance, take the above Pareto with a=2.5 and b=c=1. The variance of a full coverage loss is 2.22, while a loss excess of the deductible has variance 2.36. If losses are loaded by h x variance, an excess policyholder can buy back to full coverage for the expected value of the deductible less 0.14h. In any case, the loading method chosen will always have to be checked for its practical application to the problem at hand.

## Conclusion

The PH-transform appears to be a useful way to build in risk load, especially for severity. The power transform, i.e., the log of the Esscher transform, should also be considered, especially when it is easier to calculate. For frequency, the Esscher transform can be applied, which in practice will often just give a change in the frequency parameters.

# References

Albrecht, P (1992) A Note on a Contribution by G. Venter *ASTIN Bulletin 22* 247-54

Buhlmann, H (1980) An Economic Premium Principle *ASTIN Bulletin* 11 52-60

Delbaen, F and Haezendonck, J (1989) A Martingale Approach to Premium Calculation Principles in an Arbitrage Free Market *Insurance Mathematics & Economics* 8.4 Dec. 269-77.

Gerber, H and Shiu, E (1994) Martingale Approach to Pricing Perpetual American Options *ASTIN Bulletin* 24.2 195-220.

Venter, G (1991) Premium Calculation Implications of Reinsurance without Arbitrage *ASTIN Bulletin* 21.2 223-30.