Allocation Criteria

Gary Venter

ERM Seminar, October 2015

Allocation Criteria

- Look for business criteria that would lead to how to allocate
- Two or three that are somewhat intuitive all lead to incremental marginal allocation for a specific large class of risk measures
- If allocation is for pricing purposes – including risk-adjusted return, etc. one more criteria seems necessary
- Almost all risk measures look like means from adjusted probability distributions, so will do that

Why Allocate Capital?

- Two main reasons:
  - Pricing related – including risk-adjusted return, etc.
  - Identifying risk concentrations – tail might be driven largely by a few businesses for instance
- Different allocations with different criteria might be used for different reasons
- Basically they all come down to allocation of risk measures, so that will be the focus from here
- Total company risk measure usually less than sum of individual unit risk measures, so we are not charging the units for risk but giving them diversification credits

Criteria

- If a unit makes a small proportional change to its volume (decreasing its quota share reinsurance, perhaps) the resulting change in the company risk measure all gets allocated to that unit.
  - A practical requirement if internal competitive businesses guarding against each other
  - Consistent with economic principle of marginal-cost pricing
- If a unit with better than average risk-adjusted profit increases slightly in volume, the company risk-adjusted profit will increase.
  - Called suitability

Implications

- Two are basically equivalent
  - If additional capital from growing the unit is applied to that unit, and the unit gets higher-than-average profit on that capital, then company return on capital increases
  - Otherwise at a minimum, suitability is not guaranteed
- Both will occur iff:
  - Allocation of the risk measure is incremental marginal
    - The increase in the company risk measure due to a small proportional increase in the business unit volume, grossed up to the volume of the unit, is the allocated risk measure
  - And those incremental marginal allocations add up to the total company risk measure

So When Do They Add Up?

- Can think of the risk measure for the company as a multivariate function of the business unit random vars
- Then the allocation to the unit is the derivative of the company risk measure wrt the volume of the unit
- There is a theorem of Euler about such derivatives adding up to the function
- It happens if the function is homogeneous degree one
  - $F(ax) = aF(x)$
- Holds for almost all risk measures, but not variance: std dev, VaR, TVaR, …

What Are the Derivative-Based Allocations

- Say business unit random variable X_i part of company Y with risk measure = Q.
- TVaR allocation: $E[X_i \mid TVaR(Y) = Q]$
  - If TVaR(Y) is estimated as average of 200 largest loss simulations of the company, where every unit is part of the simulation, then average loss of X_i in those simulations is a reasonable estimate of $E[X_i \mid TVaR(Y) = Q]$
- VaR allocation: $E[X_i \mid VaR(Y) = Q]$
  - If Q is the value of Y for a single simulation, the value of X_i at that simulation is not a good estimate of $E[X_i \mid VaR(Y) = Q]$
  - Actually there are better estimates of VaR(Y) – will return to this later

For Standard Deviation

- The allocation is std dev(X_i) * correlation(X_i,Y)
  - For derivation see Venter, Major and Kreps 2006, Marginal Decomposition of Risk Measures, ASTIN Bulletin 36: 375-413
  - Basically takes derivative using L'Hopital's rule.
- Sometimes called diversified standard deviation
- Shows diversification benefit as the correlation
- Small units will have small impacts on Y so the correlation will be low – suggesting they should grow
- But as they do, correlation will increase

Risk Measure by Probability Transform

- Used mainly for risk-adjusted profit and return
- Say you are pricing a risk X by EX + 0.3*std(X)
- Transformed mean will be a value like that, > mean
- Example: exponential transform $F^*(y) = \frac{e^{kF(y)}-1}{e^k-1}$
  - Starts at 0, gets to 1, rises more slowly at first then steeper
  - $f^*(y) = \frac{e^{kF(y)}kf(y)}{e^k-1}$ so $\frac{f^*(y)}{f(y)} = \frac{e^{kF(y)}k}{e^k-1}$
  - Maybe k < 7. Numerator starts at k so ratio starts small but gets slightly greater than k.
  - Ratio called the Radon-Nikodym derivative

Advantages of Transforms

- Comes from financial pricing theory about consistency of prices of different parts of the distribution, like layers
  - Would use transform of whole distribution – same k – for pricing all layers
  - Also in simulation of many business units, would use transformed probabilities of whole company for each scenario, and use those to get target risk prices for each unit
- Extends range of available risk measures
- Makes it easier to calibrate to market prices
- In complete markets there would be a unique transform but we have to look for transforms that seem to work.

Most Risk Measures Are Actually Transforms

- VaR(0.99) is the mean when all the probability is moved to the 99th percentile
- TVaR(0.99) is the mean when losses above the 99th percentile have their probabilities multiplied by 100 and all others get zero
- Standard deviation less obvious but will get to it
- EX + kVaR(0.99), etc. is what we want to express as a transform
- Should work since linear combination of means

VaR Loading as a Transform

- Say v is the VaR at a selected percentile, with $f(v)=p$ and the loaded losses are $EX + kv$.
- Let $s = 1 – kv/(v – EX)$ and $r = kv/p(v – EX)$
- Set $f^*(v) = (r+s)p$ and $f^*(x) = sf(x)$ otherwise.
- Integral of $f^* = E^*(1) = s+rp = 1$ so this is a distribution
- $E^*(X) = sEX + rpv = EX – [kvEX – kv^2]/(v – EX) = EX + kv(v – EX)/(v – EX) = EX + kv$
- Thus the transformed mean is the VaR loaded price.
- Only risk load is for $X=v$; still layers below v get a positive load – but layers above v get a negative load
  - Not a great transform for real pricing

Pricing Layers with Transforms – VaR Example

- Risk-loaded losses calculated as mean under transformed probabilities

- Take aggregate layer with retention 0 and limit b < VaR (for now transforming aggregate probabilities)

$$E^*X = \int_0^b xf^*(x)dx + b[1 - \int_0^b f^*(x)dx]$$

- Here f* is lower than f by a factor of s. Subtracting EX:

$$-(1-s)\int_0^b xf(x)dx + (1-s)\int_0^b bf(x)dx$$

- Which is > 0. For a < b this would be less positive, so the layer from a to b would also have a positive load

- The unlimited layer with retention VaR would have price lower than expected by a factor of s

Standard Deviation as a Transform

- Consider the transform $f^*(x)/f(x) = 1 + k(X - EX)/stdX$
- Probability > 0 for $k < CV(X)$ as $f^*(0)/f(0) = 1 - k/CV$
- $E^*(X) = E[X + k(X^2 - XEX)/stdX] = EX + k[EX^2 - (EX)^2]/stdX = EX + kstdX$
- $E^*(1) = 1 + E[k(X - EX)/stdX] = 1$

Wang Transform

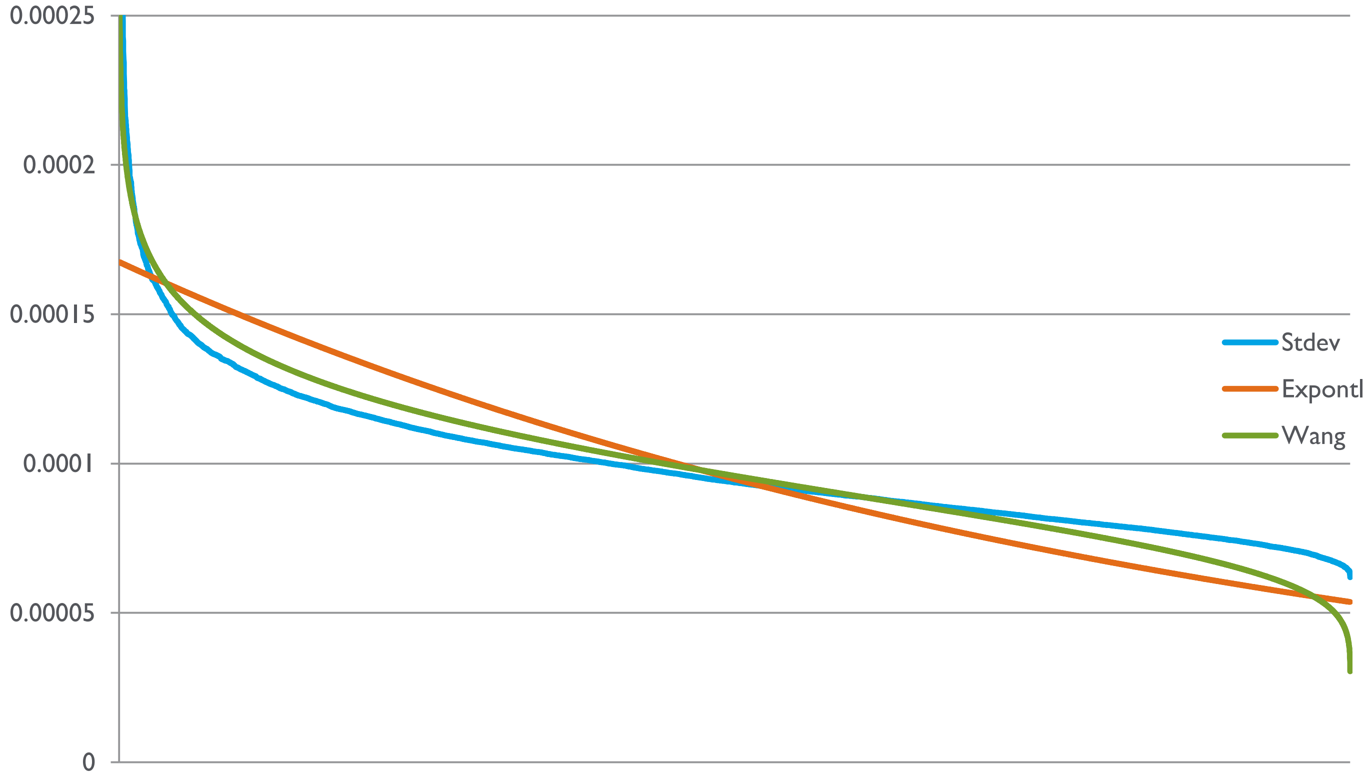
- $F^*(x) = \Phi[\Phi^{-1}(F(x)) - k]$
- Find the standard normal percentile corresponding to F(x), decrease it by k, then find the (lower) standard normal probability there. Puts more probability > x.
- Has fit well to some empirical prices

Example

- HdgFnd, Inc. has a fairly risky insurance business with 500 of annual expected losses offshore in Reunion.
- Its losses are gamma distributed with std dev 380.
- It wants to double its insurance volume and wants to start a cat business as well as high-limits casualty.
- The cat portfolio is lognormal with mean 250, CV=2.
- The casualty has 5 expected losses per year – Poisson
- Severity is Pareto $F(x) = x/(x+b)$ with limit 1000, EX=50
- Limited mean is $b*\log(1+1000/b)$ giving $b = 11.077$.

Simulation Look

- Simulated 10,000 scenarios and computed k in Exponential, Wang, and StDev transforms to give a loading of 20% overall. Transformed probabilities starting at the largest losses:



Some Numbers from Simulation

	Old	Cat	Cas	Total
Mean	495	258	248	1,000
Stdev	373	520	327	711
Corrl Total	51%	72%	44%	100%
Diversfd Stdev	192	374	145	711
Expo Transf	571	326	303	1200
Wang Transf	561	342	297	1200
Stdev Transf	549	363	289	1200

- Some simulation randomness shows in moments
- The new lines have higher correlations to the total than their % of expected losses as when they are high it is often enough to make total high as well
- Biggest pricing difference is in cat business
- Selecting a transform a judgment call based on market prices and business goals

Allocating VaR

- Doesn't work for risk-based return but maybe for identifying risk concentrations
- Just looking at makeup of 1:100 simulation doesn't work – simple simulation example
- $E[Y|F(Y) = 1 - \alpha)$
- $r(X_j) = E[X_j|F(Y) = 1 - \alpha]$
  - A 0.01% change in α leads to significant changes in LOB VaRs and the estimated allocated capital to each line.
  - Simulation is 3 gammas each with mean 1.5 and stdev 2, 1, ½.

VaR	Line 1	Line 2	Line 3	Total
98.99%	10.06	1.40	1.62	13.08
99%	8.51	2.90	1.68	13.09
99.01%	11.04	0.51	1.56	13.12

VaR	Line 1	Line 2	Line 3
98.99%	77%	11%	12%
99%	65%	22%	13%
99.01%	84%	4%	12%

Improved Allocation

- Even estimating VaR by value at one simulation is not optimal – simulation error
- Say trying to estimate 99th percentile
- Every simulation can be regarded as a Bernoulli trial for getting a number less than the 99th percentile, with a 99% probability of success
- Out of N simulations, the number of successes is binomial in N and 0.99. Say N=10,000. Then the probability that the true 99th percentile is closest to say the 9870th ranked simulation can be approximated by the average of the binomial probabilities at 9870 and 9871.
- Taking the weighted average of all the simulations by those probabilities would give another estimate of that percentile
- Then the same weighted average of the business losses at those same scenarios can be used to allocate VaR
- That binomial has stdev about 10, so 40 scenarios up & down, even with a normal distribution for weights, should be enough

VaR Continued

- Binomial method is a simplification of a method from Harrell and Davis (1982) “A new distribution-free quantile estimator,” *Biometrika* 69 3: 635-40, designed for estimation of percentiles in small samples, which interpolated the binomial with the beta distribution.
- A simpler alternative is fuzzy VaR, which takes the mean from a uniform distribution around the target simulation.
- Another interesting alternative is kernel smoothing, e.g., see Tasche (2009) “Capital allocation for credit portfolios with kernel estimators,” *Quantitative Finance* 9: 5:581-595, or Electronic copy at: http://ssrn.com/abstract=953735.

Kernel Smoothing

- Again takes a weighted average of scenarios, but instead of taking a range of scenarios around the 9900th, takes a range of values around the value v at the 9900th.
- Based on rules of thumb from the kernel smoothing business but could be changed by users
- Set $h = 1.06*stdev(X)/N^{0.2}$. For N=10,000 ~ stdev/6
- Give weight to a simulated value of y proportional to the standard normal density at $(y – v)/h$.
- Doesn't depend on what percentile you are after
- Probably sticking to y within half a stdev(X) from v is ok

Allocation of VaR in 3 Gamma Case

Method	Description	Line 1	Line 2	Line 3
VaR	Using one simulation	84%	4%	12%
Fuzzy VaR	Average 100 simulations	78%	12%	10%
Kernel VaR	Normal KDE	70%	17%	13%
Binomial VaR	Binomial	71%	16%	13%
Beta VaR	Interpolated Binomial	71%	16%	13%

Occurrence Layer Pricing by Transforms

- Allocation used transforms of aggregate probabilities but occurrence pricing needs frequency and severity distributions
- My PCAS 2004 discussion of Ruhm's paper on distribution-based pricing addresses several issues
- One is that transforms have to be applied to the ground-up distributions then the same probabilities used for every layer
- Transforming layer probabilities leads to problems that Ruhm illustrates
- Also I show methods based on Møller (2003) "Stochastic orders in dynamic reinsurance markets," *ASTIN Colloquium Papers* for simultaneously transforming frequency and severity

Frequency and Severity Transforms

- Møller has a method for building related transforms, but most come down to increasing the ground-up frequency mean by 1 / the largest reduction factor for the severity. Usually this multiplies λ by $f(0)/f^*(0)$.
- For instance, the exponential transform for HdgFnd used k = 1.1382. Since $f(0)/f^*(0) = [e^k - 1]/k = 1.86$, so this would be the increase in frequency mean if this transform were used for severity. This factor approaches unity as k goes to zero.
- This approach guarantees positive loads even for deals around small losses with reduced severity probabilities

Møller's Transforms

- Møller introduces some transforms. Using his names but my notation, with $g(y)$ for severity and λ as the frequency mean, two of these are:
- Minimum martingale measure
  - $g^*(y)/g(y) = [1 + ky/EY] / [1+k]$, which at $y=0$ is $1/[1+k]$
- Minimum entropy martingale measure
  - $g^*(y)/g(y) = (1+k)^y / E[(1+k)^Y]$, with $\lambda^* = \lambda E[(1+k)^Y]$
  - For this it helps to have k small and keep units in large chunks like millions. For instance in the HdgFnd example, the severity limit is 1000. $1.001^{1000} = 2.717$, which could be a reasonable probability increase at the limit. But that 1000 represents 1,000,000,000, which could be a strange power on 1+k.