Correlations versus Common Accident-Year and Calendar-Year Drivers for Long-Tail LoBs

A Single Composite Model for Multiple LoBs and the Economic Balance Sheet

Casualty Loss Reserve Seminar
Monday,15 September 2014
10:15 AM

Dr. Ben Zehnwirth
Managing Director
Insureware

Long tail liabilities (LOBs)

- Correlations
- Accident year drivers
- Calendar year drivers
- Seemingly Unrelated Regressions(SUR)
- Single composite model for multiple LOBs
- Risk Capital Allocation
- One year ahead statistics(CDR)
  - Variation in mean ultimates one year hence
- Economic Balance Sheet and Solvency II one year risk horizon metrics

Correlations between LOBs

- Three types of relationships
  - Process correlation
  - Parameter (trend) correlation
  - Similar trend structure implying commonality in calendar year drivers and/or accident year drivers
- Cannot measure these relationships unless LOB trend structure and process variability (volatility) modeled accurately
- Most important direction is the calendar year
- Reserve distribution correlation << Process correlation
- Highest Process correlation we have seen is 0.6!
- Highest Reserve distribution correlation is 0.2!

Correlations between LOBs

Take-Away points:

- **Most long tail LOBs exhibit zero correlation**
- **Each company is different**
- **Each LOB is different**
- **Common accident year and calendar year drivers are stronger relationships than correlations**
- **A single composite model for multiple LOBs/segments involves Seemingly Unrelated Regressions (SUR) – Zellner 1962**
- **For 40 LOBs there are 780 pairwise correlations. Most are zero. We create clusters.**

Correlation and Linearity

Correlation, linearity, normality, weighted least squares, and linear regression are closely related concepts.

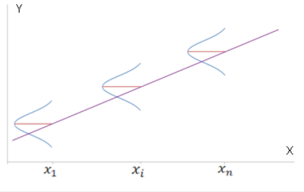
- Correlation arises naturally for two random variables that have a joint distribution that is bivariate normal.
- Two parameters (mean, standard deviation) sufficient to fully describe each individual probability distribution.
- For the joint distribution, also require additional parameter: correlation.
- If X and Y have a bivariate normal distribution, the relationship between them is linear: the mean of Y, given X, is a linear function of X i.e.

$$E(Y|X) = \alpha + \beta X$$

Correlation and Linearity

Weight = Y Height = X

For sub-populations of heights defined by $X = x_i$ the distribution of weights $Y|x_i$ is normal distribution with mean $\alpha+\beta x_i$ and variance σ^2.



Correlation and Linearity

The slope β is determined by the correlation ρ, and the standard deviations:

$$\beta = \rho\sigma_{\mathbf{Y}} / \sigma_{\mathbf{X}},$$

where $\rho = Cov(X, Y) / (\sigma_{\mathbf{X}}\sigma_{\mathbf{Y}})$.

The correlation between *Y* and *X* is zero if and only if the slope β is zero.

Also note: when *Y* and *X* have a bivariate normal distribution, the conditional variance of *Y*, given *X*, is constant i.e. not a function of *X*:

$$Var(Y|X) = \sigma^2_{Y|X}$$

Correlation and Linearity

If (Y, X) has a joint normal distribution then

$$Y \mid X = x \sim N(\alpha + \beta x, \sigma^2)$$

and

$$Var(Y) \geq Var(Y \mid X = x) = \sigma^2$$

Correlation and Linearity

This is why, in the usual linear regression model

$$Y = \alpha + \beta X + \varepsilon$$

the variance of the "error" term ε does not depend on *X*.

However, not all variables are linearly related. Suppose we have two random variables related by the equation

$$S = T^2$$

where *T* is normally distributed with mean zero and variance 1.

What is the correlation between *S* and *T* ?

Correlation and Linearity

Linear correlation is a measure of how close two random variables are to being linearly related.

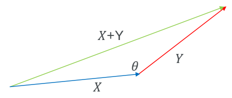
In fact, if we know that the linear correlation is +1 or -1, then there must be a deterministic linear relationship

$Y = \alpha + \beta X$ between Y and X (and vice versa).

If Y and X are linearly related, and ***f*** and ***g*** are functions, the relationship between ***f***(Y) and ***g***(X) is not necessarily linear, so we should not expect the linear correlation between ***f***(Y) and ***g***(X) to be the same as between Y and X.

(Answer to question on previous slide is zero)

The Geometry of Random Variables



$\|X\|$ is the "length" of X.
Length of a random variable = standard deviation.

Fundamental property of insurance:

$$\|X+Y\| \leq \|X\| + \|Y\|$$

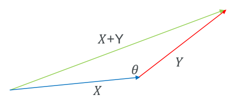
or

$$SD(X+Y) \leq SD(X) + SD(Y)$$

← The Triangle Inequality

Aggregation leads to risk diversification. Without this fact there would be no such thing as insurance.

The Geometry of Random Variables



- $\|X+Y\|^2 = \|X\|^2 + \|Y\|^2 - 2\,\|X\|\,\|Y\|\cos(\theta)$
- $Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$
- But, $\|X\|^2 = [SD(X)]^2 = Var(X)$ hence

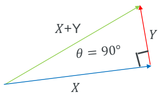
$$Cov(X,Y) = -\|X\|\,\|Y\|\cos(\theta)$$

- And since $Corr = \rho = \frac{Cov(X,Y)}{\|X\|\,\|Y\|}$ we have

$$\rho = -\cos(\theta)$$

The Geometry of Random Variables

Correlation measures the angle between two random variables.



$\rho = -\cos(90°) = 0$

If the random variables are uncorrelated ($\rho = 0$), or equivalently orthogonal ($X \perp Y$), we have Pythagoras' Theorem:

$$Var(X+Y) = Var(X) + Var(Y)$$

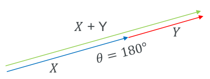
or equivalently,

$$\|X+Y\|^2 = \|X\|^2 + \|Y\|^2$$

and there is significant risk diversification

$$SD(X+Y) \ll SD(X) + SD(Y)$$

The Geometry of Random Variables



$\rho = -\cos(180°) = 1$

$$\|X+Y\| = \|X\| + \|Y\| \iff SD(X+Y) = SD(X) + SD(Y)$$

Alternatively, if $\rho = 1$ then the random variables are perfectly correlated and there is no risk diversification.

Indeed, $\rho = \pm 1$ if and only if one random variable can be written as a linear sum of the other:

$$Y = a + bX$$

If $\rho < 1$ there is risk diversification.

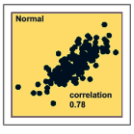
Digression: A common misconception with correlated lognormals

- Actuaries frequently need to find covariances or correlations between variables. such as when finding the variance of a sum of forecasts (for example in P&C reserving, when combining territories or lines of business, or computing the benefit from diversification).
- Correlated normal random variables are well understood.
- The usual multivariate distribution used for analysis of related normals is the multivariate normal, where correlated variables are linearly related. In this circumstance, the usual linear correlation (the *Pearson correlation*) makes sense.

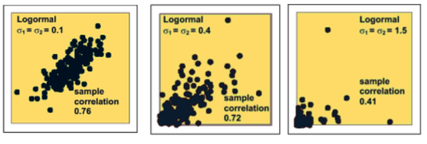
A common misconception with correlated lognormals

However, when dealing with lognormal random variables (whose logs are normally distributed), if the underlying normal variables are linearly correlated, then the correlation of lognormals changes as the variance parameters change, even though the correlation of the underlying normal does not.

A common misconception with correlated lognormals



All three lognormals below are based on normal variables with correlation 0.78, as shown left, but with different standard deviations.



A common misconception with correlated lognormals

We cannot measure the correlation on the log-scale and apply that correlation directly to the dollar scale, because the correlation is not the same on that scale.

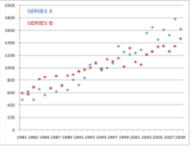
Additionally, if the relationship is linear on the log scale (the normal variables are multivariate normal) the relationship is no longer linear on the original scale, so the correlation is no longer linear correlation. The relationship between the variables in general becomes a curve**:**

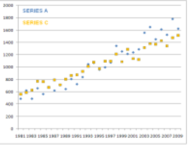
Correlation, Regression and Time Series

- Comparing Y vs time and X vs time is very different to comparing Y vs X.
- Correlations measured before and after regression can be very different.
- To assess the effective correlation between two series, must first remove trends (the predictable portion) and measure the correlation of the residuals (the random components.)

Correlation, Regression and Time Series

<u>If a time series has structure (e.g. trend) you are not measuring correlation!</u>



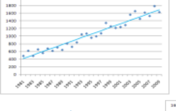


- Series A, B, and C all have a linear trend.
- B and C appear quite similar.
- The correlation between A and B is 0.91; between A and C it's 0.97

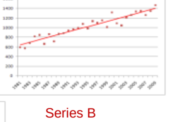
- Are A & B related? What about A & C?

De-trending the series

- Remove trends from the series.
  - In this case, using linear least-squares regression.
- Separate predictable components from the random component.

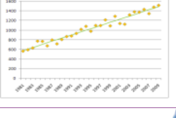


Series A

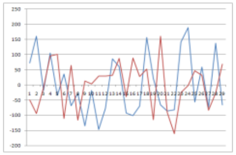


Series B

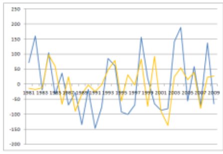
Series C



Compute the correlation of the residuals = the random component of each series



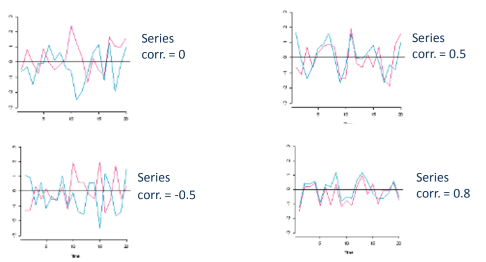
Residual or "Process" Correlation of A and B = -0.07



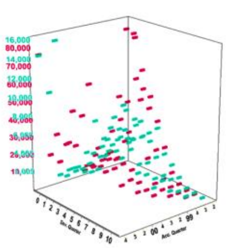
Residual or "Process" Correlation of A and C = 0.42

Conclusion: The series A and B merely share a common positive trend. There is no apparent causal or predictive relation between them. Series A and C exhibit a positive correlation. Information about the next value of C does have a significant bearing on prediction of the next value of A.

Correlation in time-series



We call the correlation of the random component (after modeling the trend structure in the three directions) of two loss development arrays: process correlation



3D plot of data

These two triangular loss arrays have process corr. = 0.9 after modelling their respective trend structures.

Cannot detect from data plot.

Correlations are in the volatility component of a model

- Two lines are (positively) correlated when their results tend to miss their target values in the same way.
- This is what should concern business planners, because it affects the unpredictable component of the forecasts.
- What is predicable when it includes common trend patterns, as in the above example, does not count towards correlation, because its effects are already incorporated into the model and forecast.

Correlations are in the volatility component of a model

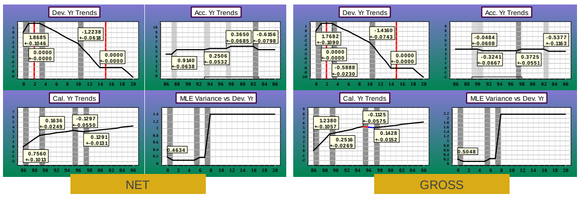
- A forecast must include a volatility measure.
- Without volatility, correlation cannot be measured. Calculating correlation requires a distribution.
- Fully-described loss distribution is ideal. But require, as a minimum, the mean and standard deviation (2nd moment) to calculate linear correlation.

Common accident year and common calendar year drivers

- Common drivers are a stronger influence than correlation.
- Not typically found outside closely related losses.
- For example, Gross versus Net of Reinsurance.
  - Net of Reinsurance is a subset of Gross so common drivers are expected.
  - Layers are subsets of ground up losses
  - Segments of the same line.
  - In this respect, detection of common drivers is as important as understanding correlations.
- The two effects must be correctly distinguished and adjusted for as management strategies of these risk components differ.

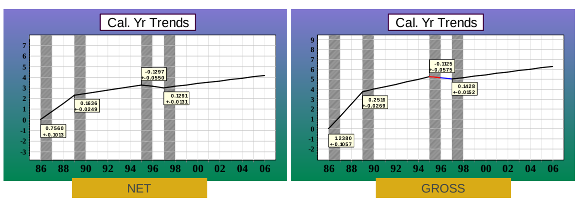
Common calendar drivers: Gross vs Net

In Gross versus Net of Reinsurance data (E&O and D&O in example), common calendar year drivers are expected to be found since Net of Reinsurance is a subset of Gross. Trends, especially calendar and accident, are closely related. The comparable models are shown below:

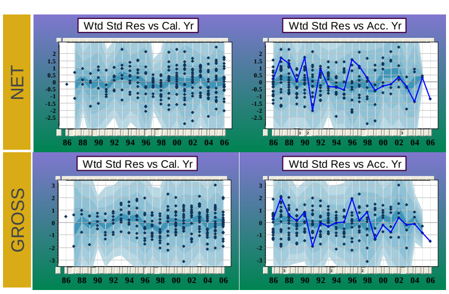


Common calendar drivers: Gross vs Net

The model trends are very similar; trend and volatility changes usually coincide. The critical trends in common are the calendar year trends (below) and accident year level changes. Common calendar year drivers are clearly visible as the trend changes occur at the same point.



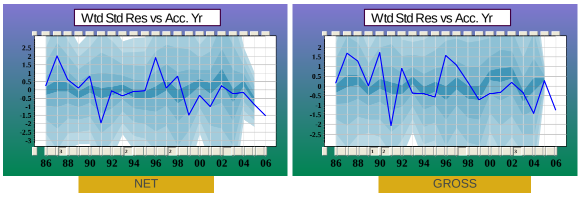
Common calendar drivers: Gross vs Net



Blue line is trace of (single) calendar year (2006) along the accident years.

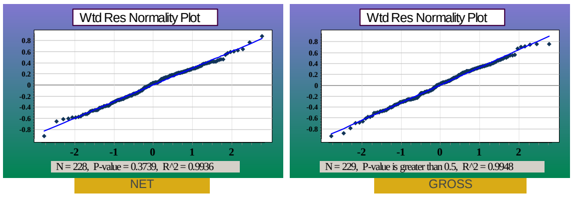
Process Correlation = 0.85

Common calendar drivers: Gross vs Net



For the model described above, the residuals by accident year traced for the last calendar year are clearly correlated; when a value in a year is low/high in one segment it is usually low/high in the other segment also at the same time.

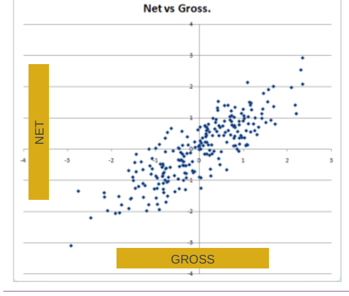
Common calendar drivers: Gross vs Net



The residuals from both lines of business are statistically indistinguishable from two normal distributions.

Thus, the process correlation can be considered the volatility correlation between two normal distributions.

Common calendar drivers: Gross vs Net



A scatter plot of the residuals, from the respective Gross and Net of Reinsurance models, exhibits a clear (linear) relationship; a correlation of 0.853.

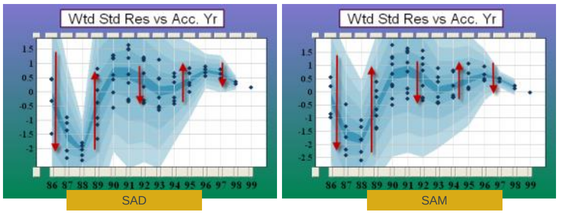
Final Correlations

Final Weighted Residual Correlations Between Datasets

	Gross PL(I)	Net PL(I)
Gross PL(I)	1	0.853
Net PL(I)	0.853	1

3 iterations were executed
Residuals correlation difference tolerance 0.010%

Common accident year drivers: SAD and SAM



A model which does not take into account the changes in accident year levels shows a marked similarity in the fluctuations of residuals in the accident direction.

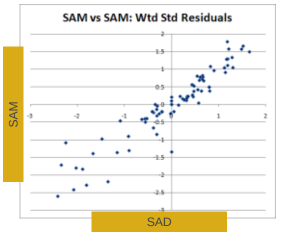
This is **not** correlation!

Common accident year drivers: SAD and SAM

Final Weighted Residual Correlations Between Datasets

	SAD:PL(I)	SAM:PL(I)
SAD:PL(I)	1	0.956
SAM:PL(I)	0.956	1

3 iterations were executed
Residuals correlation difference tolerance 0.010%

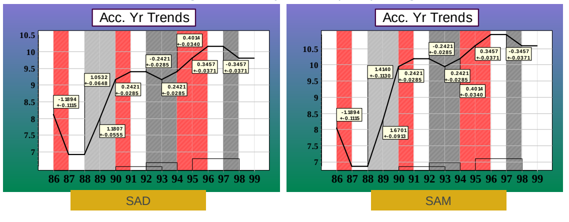


If the common accident year movements are ignored and the average accident year level fitted to both segments, then a very high spurious correlation measure of 0.96 is obtained.

The residual displays with scatterplot for SAD and SAM are shown for this model. The correlation is very high, but it is largely spurious - there are distinct changes in level across the accident years which were ignored in this model.

Common accident year drivers: SAD and SAM

Accounting for accident year level (trend) changes



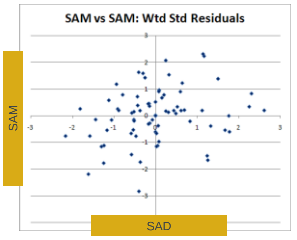
The red bars indicate common parameters between the segments. Although the calendar and development year parameters vary slightly, the accident year parameters move synchronously thus making the mean ultimates vary synchronously (but this is not correlation).

Common accident year drivers: SAD and SAM

Final Weighted Residual Correlations Between Datasets

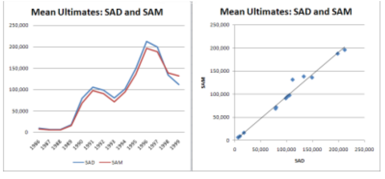
	SAD:PL(I)	SAM:PL(I)
SAD:PL(I)	1	0.249
SAM:PL(I)	0.249	1

4 iterations were executed
Residuals correlation difference tolerance 0.010%



Both sets of residuals test well for normality and have no indications of non-randomness so the process correlation (0.249) is the volatility correlation between two normal distributions.

Common accident year drivers: SAD and SAM

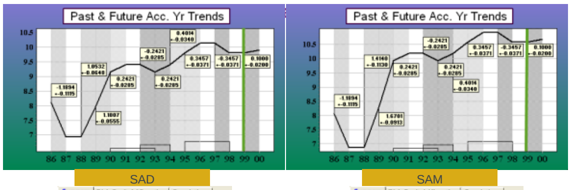


- The accident year levels moving together is a much stronger relationship than volatility correlation.
- The mean ultimates move synchronously (left) and a graph of the mean ultimates of SAM versus the mean ultimates of SAD (right) shows an almost perfect linear relationship.
- The reserve distribution correlation is only 0.086! The reserve correlation is the correlation in the losses *not* explained by the means – and therefore is the critical measure when evaluating risk diversification.

Common accident year drivers and pricing future accident years

- The linear relationship in mean ultimates is important when forecasting future underwriting (accident) years.
- If the accident year level for one segment is expected to increase by 10% $\pm$ 2%, then the other segment is also likely to increase by 10% $\pm$ 2% in the same accident year.
- The relationship between mean accident-year parameter estimates is not volatility (risk) correlation and does not indicate lack of diversification.
- The movement in means likely related to internal or external drivers. Risk exposure can be managed.
- Volatility correlation is correlation in the random component. Risk exposure here is much harder to manage as it is not able to be connected to any internal or external drivers. It is left unexplained by the model.

Common accident year drivers and pricing future accident



SAD

Summary | Risk Capital Allocation | Correlations

Combined Accident Yr Summary

Acc. Yr	Mean Outstanding	Mean Ultimate	Standard Dev.
1998	85,644	151,676	5,298
1999	122,061	111,300	7,126
2000	154,176	154,176	9,441
Total	494,382	1,542,424	20,270

1 Unit = $1
Forecast scenario: Acc. year 2000:10% + _2%?

SAM

Summary | Risk Capital Allocation | Correlations

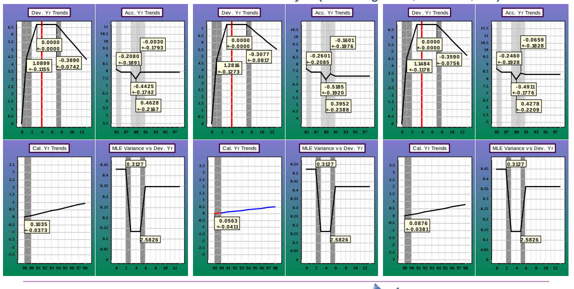
Combined Accident Yr Summary

Acc. Yr	Mean Outstanding	Mean Ultimate	Standard Dev.
1998	53,246	138,000	4,088
1999	96,842	131,073	6,029
2000	150,900	150,900	9,685
Total	408,615	1,395,346	18,646

1 Unit = $1
Forecast scenario: Acc. year 2000:10% + _2%

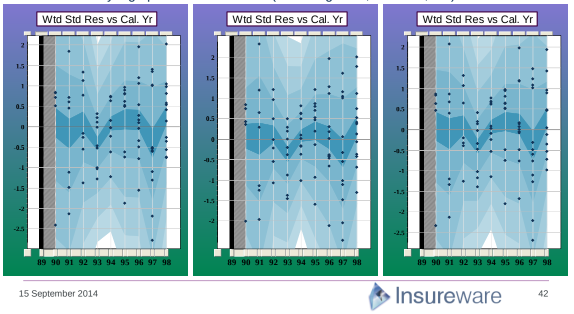
Layers Lim1M, Lim2M and 1Mxs1M; Lim2M=Lim1M+1Mxs1M

The trend structure is the same for each layer (Left to right 1M, 1Mxs1M, 2M)



Layers Lim1M, Lim2M and 1Mxs1M; Lim2M=Lim1M+1Mxs1M

Very high process correlations (Left to right 1M, 1Mxs1M, 2M)



Layers Lim1M, Lim2M and 1Mxs1M; Lim2M=Lim1M+1Mxs1M

Tables of process correlations (linear) and calendar year parameter correlations (linear)

This type of equivalent trend structure and high parameter and process correlations has not been observed for two LOBs

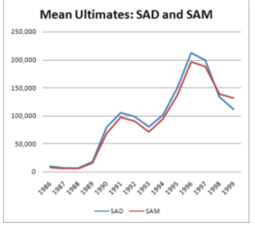
Iota Correlations

		1M:PL(I)	1Mxs1M:PL(I)	2M:PL(I)
Dataset	Period	1989~1998	1989~1998	1989~1998
1M:PL(I)	1989~1998	1	0.945646	0.992496
1Mxs1M:PL(I)	1989~1998	0.945646	1	0.977333
2M:PL(I)	1989~1998	0.992496	0.977333	1

Reserve Forecast Distributions Correlations Between Datasets (Totals)

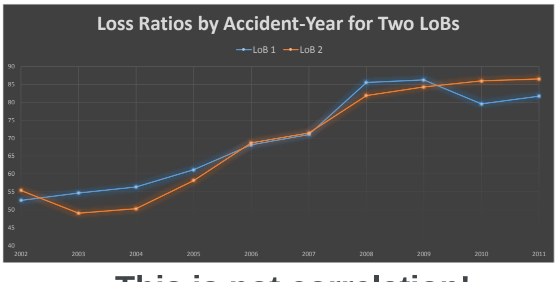
	1M:PL(I)	1Mxs1M:PL(I)	2M:PL(I)
1M:PL(I)	1	0.939207	0.991686
1Mxs1M:PL(I)	0.939207	1	0.974411
2M:PL(I)	0.991686	0.974411	1

Spurious correlation



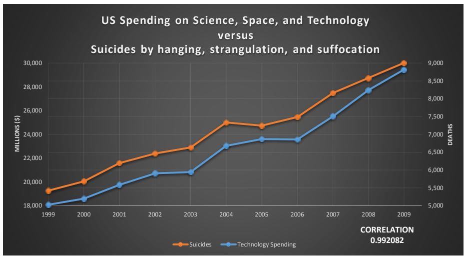
This is not correlation!

Spurious correlation



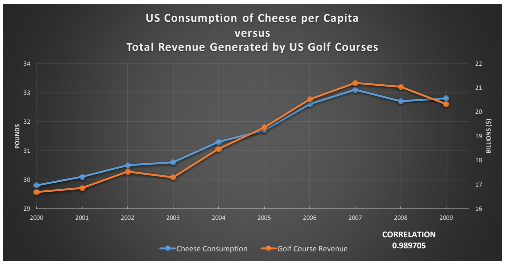
This is not correlation!

Spurious correlation



This is not correlation!

Spurious correlation

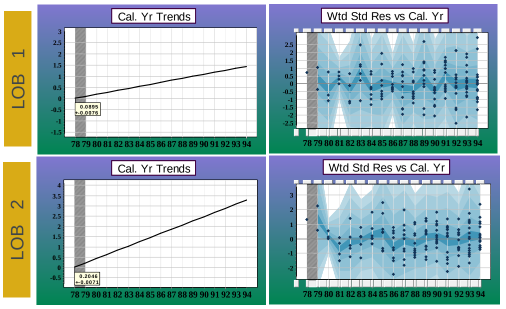


This is not correlation!

Spurious correlation

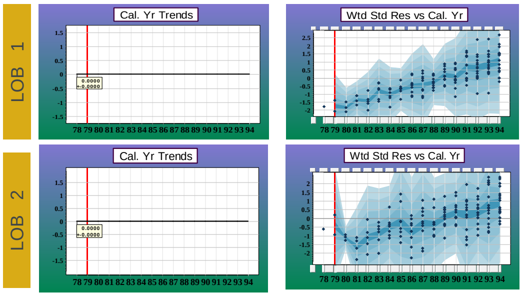
- Two LOBs are simulated independently, each with its own unique trend structure.
- One LOB has a calendar year trend of 10%, the other of 20%. Each has a -30% development year trend.
- A correct model of the underlying data process would recognise that each LOB has a separate trend for each direction and a process correlation of zero - since this is how the data were generated.
- If an incorrect model is used, one that does not describe the calendar year trends, then a spurious correlation would be detected, as an artefact of unaccounted-for structure in the data.

Spurious correlation



Correct model picks up true calendar year trend; process correlation is zero!

Spurious correlation



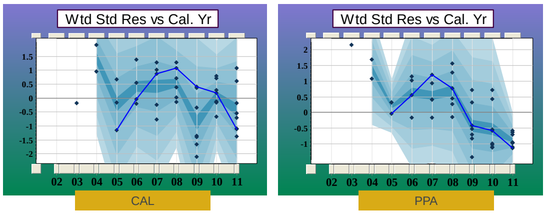
Incorrect model fails to pick up calendar trend; measures 98% correlation!
<u>But</u> this is <u>not</u> correlation since each sample is <u>not</u> random. They have structure.

Spurious correlation between Industry PPA and CAL data

- Spurious correlation is introduced by failing to detrend the data in the three directions.
- The correlation measured was spurious as there were trends in the data not described in the models.
- Once these trends were accounted for, the process correlation was statistically insignificant.

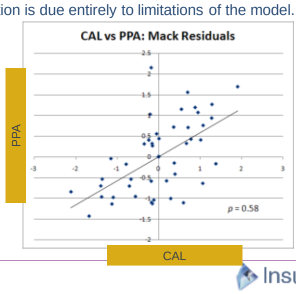
Spurious correlation between Industry PPA and CAL data due to wrong model

Paid Losses for the Industry PPA and CAL data from AM Best (2011) are modelled using the Mack method. The residuals are shown by Calendar year for CAL and PPA with the trace line for accident year 2004 highlighted.



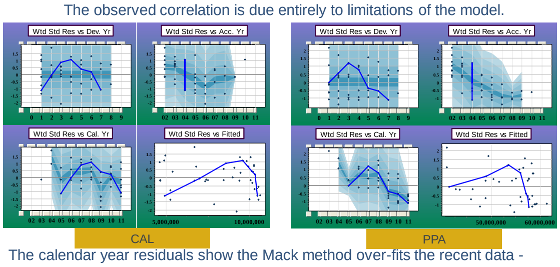
Spurious correlation between Industry PPA and CAL data

Although the residual correlation is strong the indication is misleading. The observed correlation is due entirely to limitations of the model.



Spurious correlation between Industry PPA and CAL data

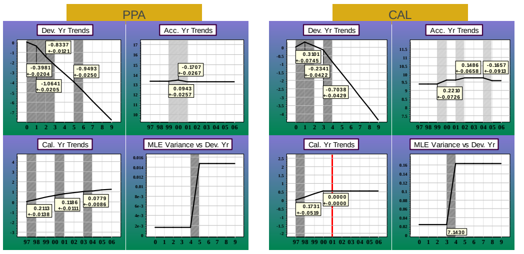
The observed correlation is due entirely to limitations of the model.



The calendar year residuals show the Mack method over-fits the recent data - producing a common negative trend in both residual displays.

Models for PPA and CAL

No LOBs have the "same" tend structure and most LOBs have zero process correlation. Consider Private Passenger Automobile and Commercial Auto Liability

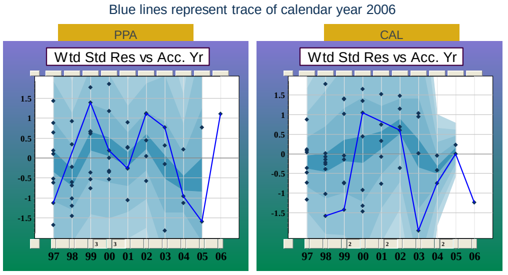


The two lines have very different trend structure and process variance!

Process correlation is zero

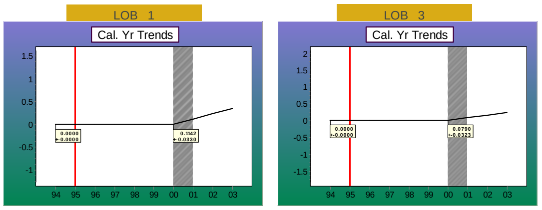
PPA and CAL have different trend structure and zero process (validation) correlation

Blue lines represent trace of calendar year 2006



Note zero process correlation.

A Tale of Two LOBs: LOB1 and LOB3

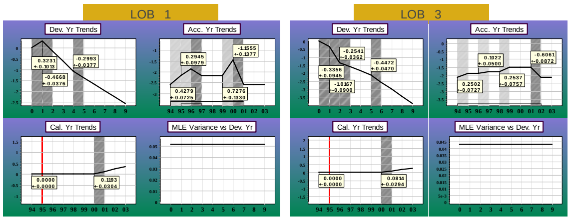


(Actually same line, different territories)

Both LOBs had a calendar year trend change in 2000

That should have been of concern!

A Tale of Two LOBs: LOB1 and LOB3



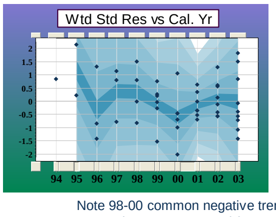
Full model display

Trends in each direction and variance of normal distributions

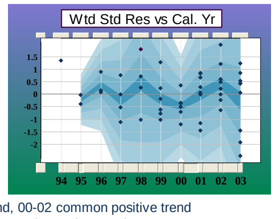
A Tale of Two LOBs: LOB1 and LOB3

Volatility correlation = Process correlation = 0.35 = Correlation in normal distributed residuals

LOB1



LOB3



Note 98-00 common negative trend, 00-02 common positive trend and 02-03 zero trend for LOB1 and negative trend LOB3.

Regression in the presence of correlation Seemingly Unrelated Regressions (SUR) – Zellner (1962)

Model displays shown above correspond to two linear models, which are described by the following equations:

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_1, \\ \mathbf{y}_2 &= \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}_2, \end{aligned} \qquad (1)$$

$$E\boldsymbol{\varepsilon}_i = \mathbf{0},\ \ i = 1, 2;\ \ E(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2^{\mathrm{T}}) = \mathrm{cov}(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2) = \mathbf{C};\ \ \ corr(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2) = \mathbf{R}$$

Without loss of sense and generality two models in (1) could be considered as one linear model:

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{pmatrix} \qquad (2)$$

Regression in the presence of correlation

Which could be rewritten as:

$$\mathbf{y} = \mathbf{X\beta} + \boldsymbol{\varepsilon}$$

For illustration of the most simple case we suppose that size of vectors *y* in models (1) are the same and equal to *n*, also we suppose that

$$E(\boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_i^T) = \mathrm{var}(\boldsymbol{\varepsilon}_i) = \mathbf{I}_n\sigma_i^2,\ i = 1, 2; \qquad \mathbf{C} = \mathbf{I}_n\sigma_{12}$$

In this case

$$\mathrm{var}(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma} = \begin{pmatrix} \mathbf{I}_n\sigma_1^2 & \mathbf{I}_n\sigma_{12} \\ \mathbf{I}_n\sigma_{12} & \mathbf{I}_n\sigma_2^2 \end{pmatrix}$$

Regression in the presence of correlation

For example, when *n* = 3

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & 0 & 0 & \sigma_{12} & 0 & 0 \\ 0 & \sigma_1^2 & 0 & 0 & \sigma_{12} & 0 \\ 0 & 0 & \sigma_1^2 & 0 & 0 & \sigma_{12} \\ \sigma_{12} & 0 & 0 & \sigma_2^2 & 0 & 0 \\ 0 & \sigma_{12} & 0 & 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_{12} & 0 & 0 & \sigma_2^2 \end{pmatrix}$$

Regression in the presence of correlation

There is a big difference between linear models in (1) and linear model (2), as in (1) we consider models separately and could not use additional information, from dependency (process correlation) of these models, what we can do in model (2). To extract this additional information we need to use proper methods to estimate vector of parameters β. The estimation

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

which derived by ordinary least square (OLS) method, does not provide any advantage, as the covariance matrix Σ does not participate in the estimations.

Only general least square (GLS) estimation

$$\tilde{\boldsymbol{\beta}} = (\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{y}$$

could help to achieve better results.

Regression in the presence of correlation

However, it is necessary immediately to underline that we do not know elements of the matrix Σ and we have to estimate them as well. So, practically, we should build iterative process of estimations

$$\tilde{\beta}^{(m)}, \ \tilde{\Sigma}^{(m)}$$

and this process will stop, when we reach estimations with satisfactory statistical properties.

The SUR β is a (credibility) weighted average β_1 and β_2

Regression in the presence of correlation

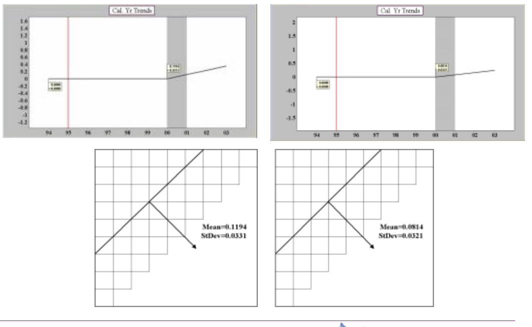
There are some cases, when model (2) provides the same results as models in (1). They are:

1. Design matrices in (1) have the same structure (they are the same or proportional to each other.)
2. Models in (1) are non-correlated, in other words

$$\sigma_{12} = 0$$

However in situation when two models in (1) have common regressors model (2) again will have advantages in spite of the identical structure of the design matrices.

Model Displays for LOB1 and LOB3 for Calendar Years



Model for individual iota parameters- they are correlated going forward

$$\hat{\iota}_1 \sim N\left(\mu_1, \quad \sigma_1^2\right); \qquad \hat{\mu}_1 = 0.1194; \quad \hat{\sigma}_1 = 0.0331$$
$$\hat{\iota}_2 \sim N\left(\mu_2, \quad \sigma_2^2\right); \qquad \hat{\mu}_2 = 0.0814; \quad \hat{\sigma}_2 = 0.0321$$

$$\begin{pmatrix} \iota_1 \\ \iota_2 \end{pmatrix} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad \hat{\boldsymbol{\mu}} = \begin{pmatrix} 0.1194 \\ 0.0814 \end{pmatrix}, \quad \hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 0.001097 & 0.000344 \\ 0.000344 & 0.001027 \end{pmatrix}$$

$$\rho = corr(\iota_1, \iota_2), \qquad \hat{\rho} = 0.359013$$

Reserve distribution correlations between two distinct LOBs - a very different story

- Highest process correlation observed between two different LOBs is about 0.6 (in our experience)
- But Reserve distribution correlation is typically lower.
- Trend structures for two LOBs typically different
- Parameter correlations low or zero
- See Private Passenger Automobile (PPA) versus Commercial Auto Liability (CAL)

Correlations and Other Relationships

There are five types of relationships.

1. Process Correlation between two sets of (random) residuals
2. Parameter Correlation
3. Same Trend Structure - Common calendar year drivers. *This is stronger than correlations.*
4. Common Accident-Year Drivers - Major implications for pricing future accident years. *This relationship is also stronger than correlations.*
5. Reserve Distribution Correlations by total, accident years and calendar years.

The optimal single composite model may also involve cross dataset parameter constraints.

Correlations and Other Relationships

- #1 induces #2.
- #3 is the 'worst' kind of relationship you can have between two LOBs
  - Very little, if any, risk diversification.
  - For future calendar year trends, the two LOBs move together. i.e. trend changes in one LOB mean trend changes in the other LOB.
- If two LOBs satisfy #3, then #1 and #2 are typically not far from 1.
- #3 – Only ever observed between layers of the same LOB, between segments of the same LOB, and between net of reinsurance and gross data (of the same LOB).

Correlations and Other Relationships

- #1, #2, #3 induce #5.
  - #5 is typically much less than #1 in the absence of #3.
- #4 results in mean ultimates by accident year moving synchronously.
  - Relationship may be close to linear- this is stronger than correlations and has implications for pricing.
  - Synchronous mean ultimates are already incorporated in the reserving model.
  - Sometimes only one or two accident years move synchronously due to a major event like Katrina. The process correlation about the new levels (trends) is usually low.
- **You cannot measure the relationship between two LOBs unless you first identify the trend structure and process variability in each LOB.**

Correlations and Other Relationships

Only in the Probabilistic Trend Family (PTF) modelling framework can you

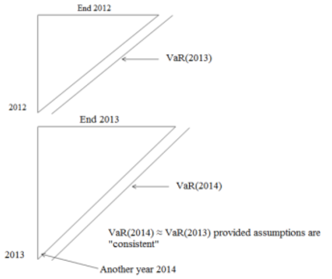
- Identify a parsimonious model that
- Separates the trend structure in the three directions from the process variability.

The data triangle (real data) is regarded as a sample path from the identified model that fits (different) normal distributions to each cell.

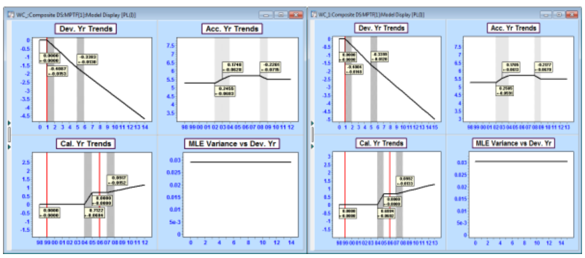
Simulated triangles from the identified good model are indistinguishable from the real data.

Updating, monitoring and the CDR

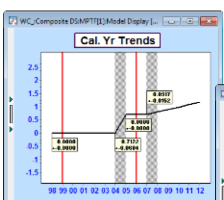
What is the CDR? "Claims development result"



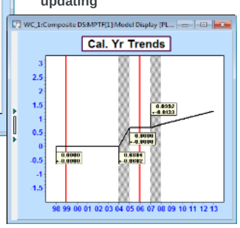
Consistent estimates of prior year ultimates and SII metrics on updating



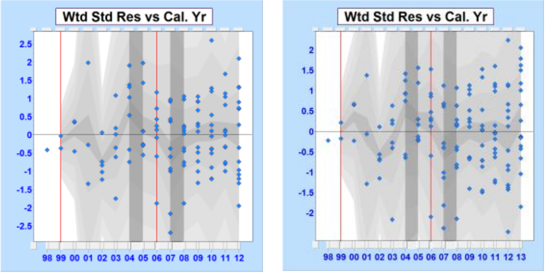
Consistent estimates of prior year ultimates and SII metrics updating



Calendar year trend has not changed significantly on updating



Consistent estimates of prior year ultimates and SII metrics updating



Consistent Estimates of prior year ultimates on updating

Original Forecast

Acc. Yr	Mean	
	Outstanding	Ultimate
1998	0	4,302,916
1999	100,508	5,822,005
2000	251,137	6,713,358
2001	487,526	9,223,056
2002	790,374	10,494,535
2003	1,518,061	16,457,237
2004	2,836,326	24,336,373
2005	4,416,971	35,141,812
2006	6,913,226	42,560,786
2007	10,939,532	48,364,178
2008	15,873,848	56,111,850
2009	17,989,038	52,985,955
2010	27,146,680	60,327,875
2011	40,184,672	66,475,390
2012	50,057,940	60,673,332
Total	179,505,841	499,990,660

8% projected CY trend

On Updating

Acc. Yr	Mean	
	Outstanding	Ultimate
1998	0	4,355,867
1999	75,545	5,897,616
2000	182,316	6,779,558
2001	342,967	9,364,055
2002	540,509	10,621,307
2003	1,025,994	16,554,493
2004	1,957,243	24,371,682
2005	2,987,693	35,500,697
2006	4,594,073	42,386,716
2007	7,156,370	47,387,224
2008	10,240,367	54,223,948
2009	11,283,994	53,265,077
2010	16,672,321	58,883,329
2011	24,238,376	60,953,881
2012	29,710,400	52,768,539
2013	43,198,738	54,117,218
Total	154,206,907	537,431,208

0% projected CY trend

Acc. Yr	Mean	
	Outstanding	Ultimate
1998	0	4,355,867
1999	81,810	5,903,881
2000	204,288	6,801,530
2001	396,169	9,417,257
2002	641,357	10,722,155
2003	1,246,396	16,774,895
2004	2,426,713	24,841,152
2005	3,769,925	36,282,929
2006	5,884,361	43,677,004
2007	9,283,300	49,514,154
2008	13,426,282	57,409,863
2009	14,925,633	56,906,716
2010	22,144,416	64,355,424
2011	32,231,974	68,947,479
2012	39,488,963	62,547,102
2013	57,378,110	68,296,590
Total	203,529,696	586,753,997

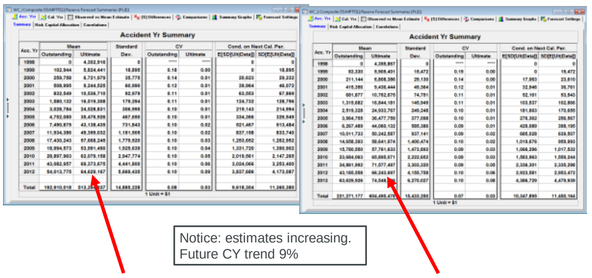
8% projected CY trend

Acc. Yr	Mean	
	Outstanding	Ultimate
1998	0	4,355,867
1999	92,271	5,914,342
2000	243,318	6,840,560
2001	496,438	9,517,526
2002	842,462	10,923,260
2003	1,710,138	17,238,637
2004	3,465,971	25,880,410
2005	5,586,472	38,099,476
2006	9,018,576	46,811,219
2007	14,671,775	54,902,629
2008	21,820,356	65,803,937
2009	24,874,620	66,855,703
2010	37,578,030	79,789,038
2011	55,389,045	92,104,550
2012	68,440,571	91,498,710
2013	100,105,162	111,023,642
Total	344,335,206	727,559,507

20% projected CY trend

For forecast mean ultimates to be consistent you require a consistent model and consistent assumptions about the future.

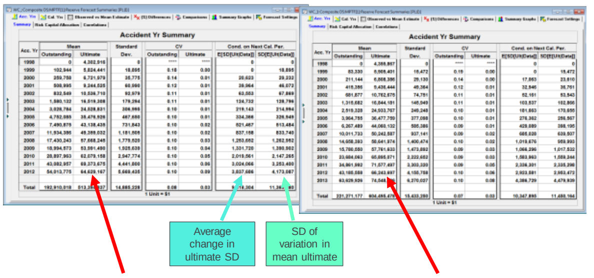
Consistent Estimates of prior year ultimates on updating



Notice: estimates increasing.
Future CY trend 9%

At end 2012, Mean ultimate distribution 2012 = 64.6 and SD of ultimate distribution = 5.7
At end 2013, Mean ultimate distribution 2012 = 66.2 and SD of ultimate distribution = 4.2

Consistent Estimates of prior year ultimates on updating

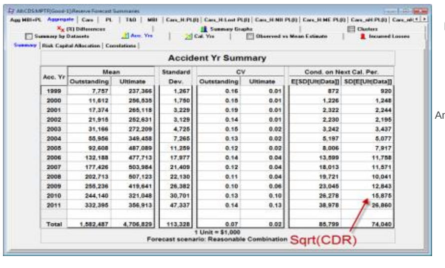


At end 2012, Mean ultimate distribution 2012 = 64.6 and SD of ultimate distribution = 5.7
At end 2013, Mean ultimate distribution 2012 = 66.2 and SD of ultimate distribution = 4.2

Peter England's "CDR" is simply Var[E[Ult.|CY1]]

When do estimates of prior year ultimates stay consistent on updating (next valuation period) ?

With identified optimal parametric distribution models that are tested from the data, it is relatively straightforward to compute the CDR.



Accident Yr Summary

Acc. Yr	Mean Outstanding	Mean Ultimate	Standard Dev.	CV Outstanding	CV Ultimate	Cond. on Next Cal. Per. E[SD[Ult|Data]]	Cond. on Next Cal. Per. SD[E[Ult|Data]]
1999	7,757	237,366	1,267	0.16	0.01	872	920
2000	11,612	256,535	1,750	0.15	0.01	1,226	1,248
2001	17,374	265,118	3,229	0.19	0.01	2,322	2,244
2002	21,915	252,631	3,129	0.14	0.01	2,230	2,195
2003	31,166	272,209	4,725	0.15	0.02	3,242	3,437
2004	55,956	349,458	7,265	0.13	0.02	5,197	5,077
2005	92,608	487,089	11,259	0.12	0.02	8,006	7,917
2006	132,188	477,713	17,977	0.14	0.04	13,599	11,758
2007	177,426	503,984	21,409	0.12	0.04	18,013	11,571
2008	202,713	507,123	22,130	0.11	0.04	19,721	10,041
2009	256,236	419,641	26,382	0.10	0.06	23,045	12,843
2010	244,140	321,048	30,701	0.13	0.10	26,278	15,875
2011	332,395	356,913	47,337	0.14	0.13	38,978	26,860
Total	1,582,487	4,706,829	113,328	0.07	0.02	85,799	74,040

1 Unit = $1,000
Forecast scenario: Reasonable Combination

Sqrt(CDR)

Note Pythagoras's theorem:

Var[Ult.] = E[Var[Ult.|CY1]] + Var[E[Ult.|CY1]]

Analogous to One-Way ANOVA

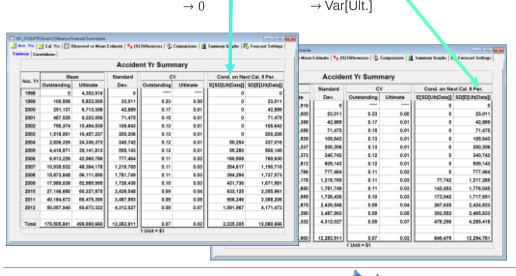
Total SS = Within Group SS + Between Group SS

Var[E[Ult.|CY_k]]

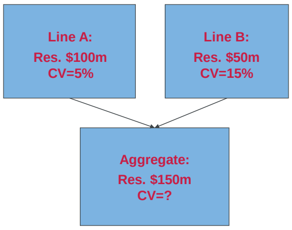
Conditioning on more future years…

Var[Ult.] = E[Var[Ult.|$CY_1, \ldots, CY_k$]] + Var[E[Ult.| $CY_1, \ldots, CY_k$]]

$\rightarrow 0$ $\rightarrow$ Var[Ult.]



Risk Capital Allocation



Assume Risk Capital at 98th percentile = 2 Standard Deviations

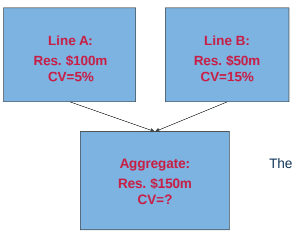
Risk Capital for Line A = $10m

Risk Capital for Line B = $15m

Aggregate Risk Capital (ARC) = $25m ?

15 September 2014 Insureware 82

Risk Capital Allocation



Assume Risk Capital at 98th percentile = 2 Standard Deviations

Risk Capital for Line A = $10m

Risk Capital for Line B = $15m

Aggregate Risk Capital (ARC) = $25m ?

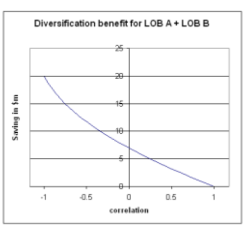
The answer depends on the correlation.

If Corr = +1.0, ARC = $25m

If Corr = 0.0 ARC = $18m

15 September 2014 Insureware 83

Risk Capital Allocation: Diversification benefit

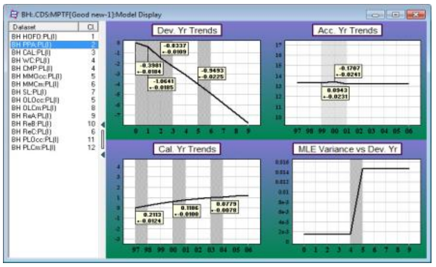


Benefit =

Sum of individual risk capital assessments – aggregate risk capital assessment from joint distribution of the two (correlated) lines.

15 September 2014 Insureware 84

A single composite model for multiple LOBs



A single composite model for multiple LOBs.

LOBs are in same cluster if significantly correlated

Correlations

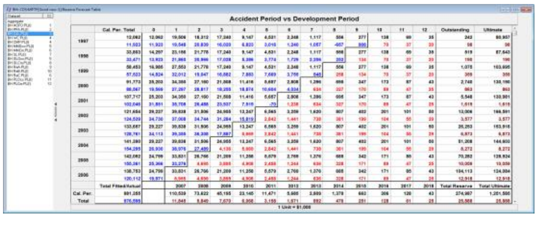
Final Weighted Residual Correlations Between Datasets

	Clusters		Datasets		Correlations	
	Select	Action	Action	Name		
1	SEL			BH HOFO:PL(I)	1	
2			Add to SEL	BH PPA:PL(I)	1	
3			Add to SEL	BH CAL:PL(I)	1	
4		Combine	Add to SEL	BH WC:PL(I)	1	0.357
			Add to SEL	BH CMP:PL(I)	0.357	1
5		Combine	Add to SEL	BH MMOcc:PL(I)	1	0.453
			Add to SEL	BH OLOcc:PL(I)	0.453	1
6		Combine	Add to SEL	BH MMCm:PL(I)	1	0.354
			Add to SEL	BH ReC:PL(I)	0.354	1

Clusters have been set
9 iterations were executed
Residuals correlation difference tolerance 0.010%

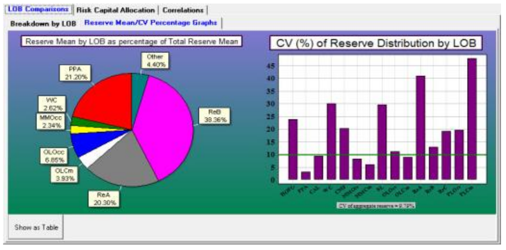
A single composite model for multiple LOBs

Projected lognormals for each cell and their correlations.
Blue is observed.
Black is fitted mean of lognormal.
Red is standard deviation of fitted lognormal.
Burgundy is standard deviation.



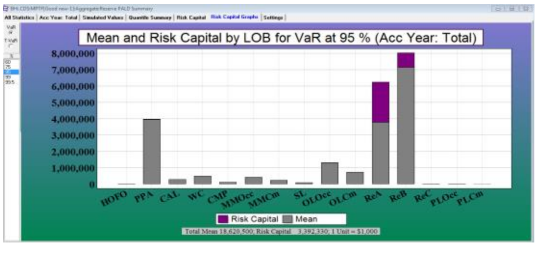
A single composite model for multiple LOBs

Note risk diversification due to lack of process correlations

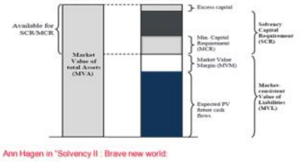


A single composite model for multiple LOBs

Risk capital by LOB for V@R at 95%



Solvency II – Economic Balance Sheet



Solvency II aims to establish a solvency regime that is better matched to the <u>true</u> risks of an insurance company

Ann Hagen in "Solvency II : Brave new world:

"Doing the job

Under Solvency II, the way that work is carried out will change. For example, Solvency II is likely to require different actuarial techniques from the ones currently used. Technical provisions will be estimated as a probability-weighted average of expected future cash flows, taking into account the time-value of money and including a risk margin. Many of us are estimating claims reserves using traditional deterministic actuarial techniques, primarily relying on incurred claims data. Under Solvency II, not only will we need to discount these reserve estimates, requiring projected payment patterns, we will also need to demonstrate a deep understanding of the uncertainty of those reserves. We will additionally be required to apply the same approach to evaluating unexpired risk liabilities currently allowed for in the unearned premium reserves."

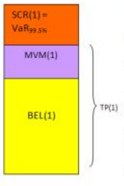
Solvency II one-year risk horizon:
* satisfies 3 conditions
* decomposing the directives
* What are the basic elements?

- Risk Capital is raised at the beginning of each year and any unused capital is released at the end of the year;
- The analyses are conditional on the first (next) calendar year being in distress (99.5%);
- At the end of the first year in distress, the balance sheet can be "restored" in such away that the company has sufficient technical provisions (fair value of liabilities) to continue business or to transfer the liabilities to another risk bearing entity.

An important consideration is that **funqibility** by calendar year is only in the forward direction

Risk Capital – One Year risk Horizon

Simplest Case: Only One Year Runoff



L_1 = projected losses for the year. This is a random variable.

$BEL(1) = E(L_1)/(1+d)^{0.5}$ Where d = interest rate. Losses are paid uniformly through year, so we discount for half a year.

$SCR(1) = VaR_{99.5\%}(L_1)$, i.e. $\Pr(L_1 \le E(L_1) + SCR(1)) = 0.995$

MVM(1) is the cost incurred in having risk fund of *SCR(1)* available for the year. It is paid to capital provider at end of year and so is discounted by a full year.

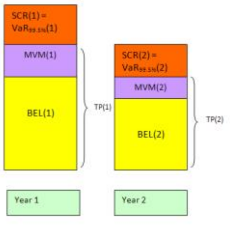
$MVM(1) = \frac{SCR(1)*s}{(1+d)}$, if the interest on the risk fund is paid directly to capital provider, or $MVM(1) = \frac{SCR(1)*(s+d)}{(1+d)}$, otherwise.

$TP(1) = BEL(1) + MVM(1)$. This is the Technical Provision and must be held in company own funds.

We will also let, *PV(k;d)*, or *PV(k)* be used to abbreviate the Present Value factor $1/(1+d)^k$

Risk Capital – One Year risk Horizon

Next Simplest Case: Two Year runoff, No correlation



$$BEL(1) = E(L_1) * PV(0.5)$$

$$BEL(2) = E(L_2) * PV(1.5)$$

$$MVM(1) = VaR_{99.5\%}(1) * s * PV(1)$$

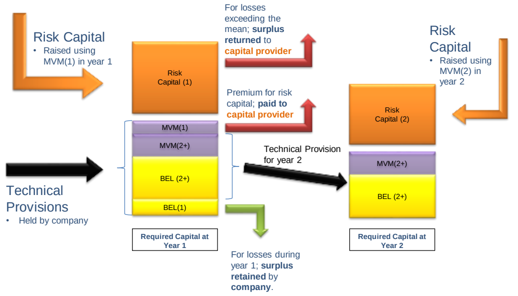
$$MVM(2) = VaR_{99.5\%}(2) * s * PV(2)$$

The Technical Provision (TP) at inception is the sum of the individual year TPs:

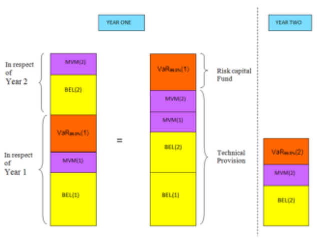
$$TP = TP(1) + TP(2)$$

This amount needs to be available in company own funds to ensure that losses can be met up to a 99.5% or 1/200 risk level in each year. Aggregate losses up to the value of the mean are met out of *BEL* funds, excess losses are met from the *SCR* fund, access to which is financed by *MVM*.

Capital flow: Uncorrelated future calendar years



Risk Capital – One Year risk Horizon



Two-year picture of accounts: In year 1 we require reserves to meet paid loss liabilities for years 1 and 2 and we also need to able to fund the cost of access to the risk capital funds for years 1 and 2, however we only need access to the year 1 risk fund. When year 2 begins our accounts reset, since any cost over-runs from year 1 were paid out of the risk fund and do not degrade our prepared reserves for year 2. Provided the loss over-run is below RC(1) = $VaR_{99.5}(L1)$.

Risk Capital – One Year risk Horizon

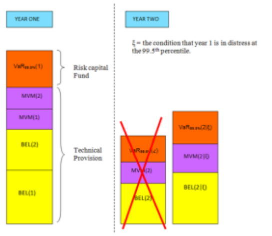
•This is fine, except for one thing:
What if the distribution for the losses in year 2 has changed conditional on the losses in year one?

•Simply put, the previous picture assumes there is no correlation between the distributions for years 1 and 2. In other words, whatever the outcome observed after year 1 we are going to remain fixed on our previous course, full steam ahead

Typically calendar year distributions are positively correlated.

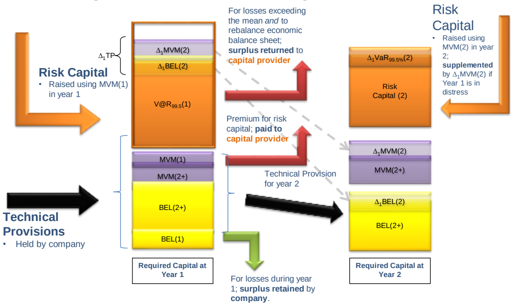
The correlations are driven by parameter uncertainty.

Risk Capital – One Year risk Horizon

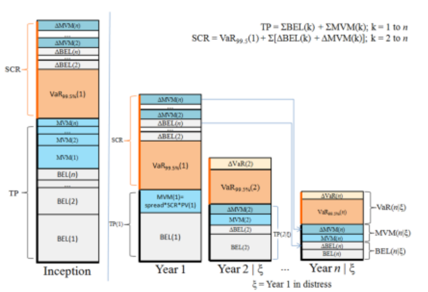


If year 1 is in distress at the 99.5th percentile, then our risk fund carries us over into year 2, but the conditional distributions are now different. Year 2 now must be re-evaluated in the light of conditional distributions and these increase the size of the BEL and the MVM, the cost of holding the risk fund. We need to include these adjustments in the year 1 risk fund

Capital flow: Two-year runoff with first year in distress



N-year run-off (Correlated)



Two-year runoff with first year in distress

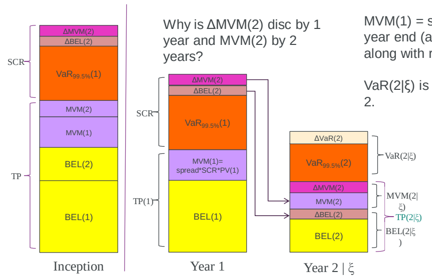
Let ξ = Year 1 in distress

Why is ΔMVM(2) disc by 1 year and MVM(2) by 2 years?

VaR(1) is consumed.

MVM(1) = spread*SCR at year end (and returned along with risk free rate).

VaR(2|ξ) is raised in year 2.



Two-year runoff with first year in distress

- There is sufficient risk capital SCR and Fair Value to withstand a distressed first year at 99.5% confidence and restore Fair Value at beginning of the second year.
- An important consideration is that fungibility by calendar year is only in the forward direction.

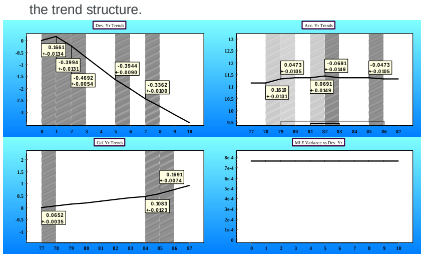
Consistent metrics on updating from year to year- under what conditions?

See also E&Y GNAIE paper (2007)

"Market Value Margins for Insurance Liabilities in Financial Reporting and Solvency Applications , October 1, 2007"

Dataset ABC: The PTF model

The optimal PTF identified model. Note the model fits a normal distribution to each cell. The means are related via the trend structure.

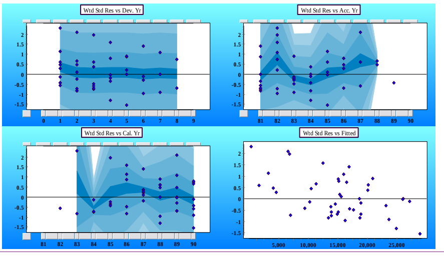


Note major calendar year trend shift

IL(C) Data

Mack (=volume weighted average) weighted standardized residuals

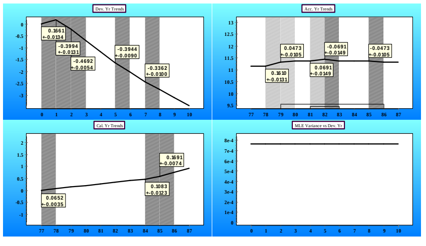
- Note trend in residuals versus fitted values (bottom right)



103

Dataset ABC: The PTF model

The optimal PTF identified model. Note the model fits a normal distribution to each cell. The means are related via the trend structure.

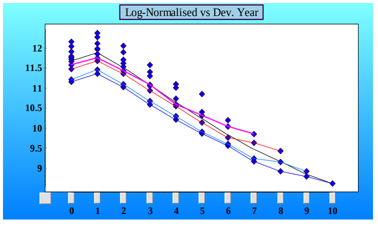


Note major calendar year trend shift.

104

Dataset ABC

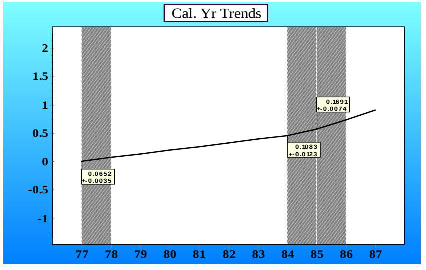
- As you move down the accident years the “kick-up” is one development period earlier
- Real data satisfies axiomatic trend properties.



105

Dataet ABC
PTF-Calendar Year Trends

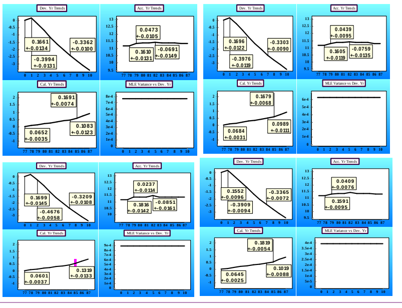
Have control on future assumptions



106

Dataset ABC

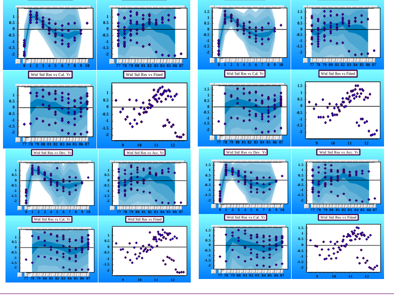
Three simulated triangles from the fitted model, and the real data triangle? Which is real data?



107

Dataset ABC

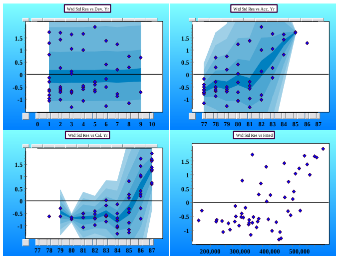
Three simulated, one real. Residuals of fitting only one parameter in each direction. Which is the real data? Simulated triangles have the same statistical features as the real data! We will use Bootstrap technique later to do same thing.



108

Dataset ABC- Wtd Standardized Residuals of Mack method (CL link ratios)

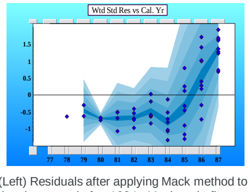
It is impossible for any link ratio method including Mack (=CL ratios) to capture and describe trends in any direction, let alone the calendar years.



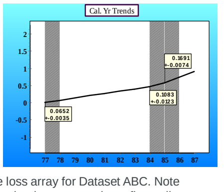
Dataset ABC

ELRF- Mack (volume weighted average link ratios) Residuals versus calendar year. Cannot capture calendar year trend structure. No control on assumptions going forward either, and averager calendar year trend captured cannot be discerned.

Mack Residuals



Calendar Year trends in incrementals



(Left) Residuals after applying Mack method to the loss array for Dataset ABC. Note the sharp trend after 1984. <u>Mack underfits recent calendar years and overfits earlier</u> years.
(Right) Probability Trend Family model picks up the change in trend structure in this direction, the other two directions and the volatility.

Dataset ABC

Removing the three calendar year trends. (setting the trend to zero for all calendar years in the PTF modelling framework)

Looks a bit like the Mack residuals (but on a log scale)