Modeling with the Multivariate Probabilistic Trend Family

Glen Barnett, Ph.D. and Ben Zehnwirth, Ph.D., AIA, AIAA[1]

1 Professorial Visiting Fellow, School of Actuarial Studies, Australian School of Business, University of NSW

Abstract: This paper motivates the benefits of modeling trends, volatility and correlations through a study of real data triangles.

We show that a model that is demonstrably unable to forecast the recent past of the historical triangle cannot be expected to tell us anything useful about the future of the same process. Naturally, the same basic model fitted by applying a more sophisticated tool will suffer the same fate. The use of GLMs, bootstrapping, or Bayesian statistics cannot avoid the basic defects of traditional methods.

With traditional techniques the parameters (e.g. age-to-age factors) are a function of the data. By contrast, in the Probabilistic Trend Family (PTF) modeling framework the model design as well as the parameters are a function of the data. We illustrate PTF modeling (e.g., Barnett and Zehnwirth, 2000) on a variety of real triangles.

The PTF modeling framework is extended to the simultaneous modeling of multiple triangles. The multivariate modeling framework (MPTF), apart from describing the volatility in each triangle also describes correlations between them in two different ways.

The MPTF modeling framework can be used in a number of innovative ways yielding useful information about the risk characteristics of the business. There are important implications for economic capital calculations and optimal retention. In order to compute economic capital for reserve risk and underwriting risk the correlations between lines of business need to be known. To assess the correlations accurately (whether from related trends or correlated errors), a model for each line that describes the trend structure and the volatility about the trend structure needs first to be identified.

SECTION 1: RATIO MODELS AND THE CHAIN LADDER

The intuitive basis of ratio models is simply that it is expected that if a known cumulative paid loss (or an incurred loss) in one period is high, that the corresponding figure for the next period will also tend to be high, and if it is low, it will subsequently tend to be low. The *ratio assumption* holds that the next value is expected to be in direct proportion to the current known figure.

In what follows, we will drop the parenthetical reference to incurred losses, but the subsequent discussion will (with possibly minor but obvious alterations) continue to apply.

Notation

We label the accident years in the triangle $i = 1, \ldots n$, and the calendar years similarly, $t = 1, \ldots n$. For ease of notation, development year is taken as the delay between accident and payment year ($j = t - i$), and hence $j = 0, 1, 2, \ldots, n-1$. Let the incremental amount be p_{ij} , and let the cumulatives total paid or incurred to date in an accident year be c_{ij}. Further, let y_{ij} be the logarithms of incrementals (possibly after adjusting for economic inflation or exposures).

Ratios

If C_{ij} is the figure for accident period i, development period j then an obvious way to record that intuitive expectation is

$$\mathrm{E}(C_{ij} \mid C_{i,j-1} = c) = \beta_j c . \tag{1}$$

There are a variety of ways to estimate the ratio parameters, β_j. Different estimators will correspond to different implicit assumptions about the remainder of the model.

On a worldwide basis, ratio methods continue to be very popular, and the chain ladder is undoubtedly the most popular of the ratio methods in use. The standard chain ladder estimators correspond to assuming that the variance of the C_{ij} is proportional to the mean (in the sense that the estimators and subsequent forecasts are optimal under that assumption; otherwise better estimators exist).

These assumptions underly chain ladder forecasts. If we consider a particular accident year, i, and j-1 is the last observed development in that year, the chain ladder forecast of C_{ij} (the first future cumulative payment in that year) follows equation (1).

There are a number of models that reproduce the chain ladder. The approach of Mack (1993) and Murphy (1994) reproduces chain ladder forecasts by combining equation (1) with the variance-proportional-to-mean assumption. It is possible to derive predictive means and variances from that.

The other widely used model that reproduces the chain ladder forecasts is the (quasi-) Poisson model with log-link and linear predictor corresponding to a two-way main-effects ANOVA. This Poisson model was introduced to actuaries by Hachemeister and Stanard (1975). The approach became popular in the 1990s after the paper by Renshaw and Verrall (1994). This model is now popularly referred to as the overdispersed Poisson model (ODP) – however the term "overdispersed" can be incorrect because the scale parameter ϕ, may (with some choices of scale, such as may occur if the data are measured in millions or billions of dollars) easily be less than 1, and therefore underdispersed. The term quasi-Poisson is more appropriate, since it applies irrespective of the size of the scale parameter.

Within the data, this model has a different structure for the mean, which corresponds essentially to the assumption:

$$\mathrm{E}(C_{ij}) = \beta_j \, \mathrm{E}(C_{i,j-1}) . \tag{2}$$

However, the *forecasting function* of this model has the same structure as (1) – as indeed it must in order to reproduce the chain ladder forecasts. For example, just as with the Mack model, for a given accident year, i, and j-1 is the last observed development, the chain ladder forecast of C_{ij} follows equation (1).

SECTION 2: ASSESSING THE CHAIN LADDER ASSUMPTIONS

2.1 Chain ladder assumptions and diagnostics

Since the primary interest for actuaries lies in the forecasting performance of these models (rather than the estimates within the data), it is crucial to test that the structure described by model (1) for ratio models in general; it is still necessary to assess within-data fit (such as via residual plots), but it is also necessary to assess the appropriateness of the prediction equation if the two differ. This is discussed in detail in Barnett, Zehnwirth, and Odell (2008), where one-step-ahead prediction errors are used to discover inadequacies of the *predictive performance* of the quasi-Poisson GLM, and it is shown that ordinary residuals do not indicate problems with the predictive properties of this model.

For present purposes, it suffices to approximate the predictive performance of the quasi-Poisson GLM by looking at displays of the residuals from the Mack model as an adjunct to the residual analysis of the GLM fit. Mack residuals are a good approximation because the one-step-ahead prediction errors and the Mack residuals are based on the same estimates – but in the case of the quasi-Poisson GLM, on fewer observations (the Mack residuals use data from all available accident years for its residuals, while the GLM one-step-ahead predictions don't use data from accident years later than the observation under consideration). In practice, simply using the Mack residuals is an effective way to assess the predictive performance of the two-way quasi-Poisson GLM with log-link.

Consequently, for the present paper, unless otherwise noted, we will use residuals from a Mack-style model (modified as indicated below) as an indicator of predictive performance of both of the popular chain-ladder-reproducing models.

ELRF models

In Barnett and Zehnwirth (2000), we expanded the framework of Murphy (who added intercepts) to construct diagnostic tests of ratio models, including the chain ladder, constructing the Extended Link Ratio Family of models (*ELRF*).

We then analyzed a number of real data triangles using standard regression diagnostics in the ELRF framework and showed that the assumptions of ratio models failed for all of the data sets analyzed. In this paper, we have carried out a small study using such diagnostics on randomly chosen triangles (in section 3) to assess the suitability of the chain ladder for a variety of different triangles.

2.2 Further issues with the chain ladder

Barnett, Zehnwirth, and Dubossarsky (2005) describe the incremental chain ladder forecasts are the same whether cumulation and ratios are done across development years or down accident years. This property – that development and accident years are interchangeable – has some worrying implications for all chain ladder-reproducing models.

These include:

(i) The fact that the accident and development directions are not distinguished by the model, despite the fact that we know them to behave quite differently;

(ii) being fully parameterized (implicit or explicit) in both those directions – in spite of the fact that the development years and accident years may both be described with only a few parameters;

(iii) Both directions exhibit clear relationships within themselves (though of different kinds), but the model doesn't attempt to utilize this information.

A second issue occurs in the case of the quasi-Poisson GLM, which we don't believe has been noted elsewhere. It is possible to add calendar-period parameters to the quasi-Poisson model, and it has been suggested that this would be suitable for dealing with changing inflation. However, this is not the case, as can be seen by a simple argument.

Imagine that the portfolio has been stable for many years, the shape of the runoff has remained unchanged, and that superimposed inflation has been zero. Then, for the most recent calendar year, superimposed inflation at 10% occurs. Clearly, the mean for each observation is 10% higher than the corresponding observation a year earlier, and – since $\mathrm{Var}(cX) = c^2\ \mathrm{Var}(X)$ – the standard deviation is also 10% higher. Yet if we model the calendar period changes with parameters in the quasi-Poisson model, the model asserts that the standard deviation will just under 5% higher. The model description of the variance is incorrect. Consequently, prediction intervals – including bootstrap prediction intervals – based on this model will be too narrow.

Note that if the inflation is random, rather than constant, the induced variance will increase still faster.

The bootstrap and Bayesian methods

Refinements such as the bootstrap and the use of Bayesian methods don't alter the adequacy or lack or adequacy of the mean and variance components of a given model – if the mean and variance of the data are not well described without them, the mean and variance will not be will described with them (unless those components of the model itself are altered).

SECTION 3: THE CHAIN LADDER EXPERIMENT

Design

A random sample of 25 complete triangles from Schedule P data was taken. Either the Paid or Incurred triangle was used (60% chance of selecting paid, 40% chance of selecting incurred).

For each such triangle, the suitability of the ratio (chain ladder-like) assumptions were tested according to a number of criteria

Criteria

There were four primary diagnostic criteria on which the chain ladder assumptions were assessed.

1) Do the lines relating incremental to previous cumulative need an intercept?

2) Are the correlations between incremental and previous cumulative "small?"

3) Do the residuals exhibit substantial changing trends against calendar periods?

4) Do the residuals have trends against fitted values?

If the answer to any of these questions is a clear yes, the chain ladder model does not hold.

In addition, there was a fifth criterion:

5) After removing a linear trend down the accident periods, are remaining linear correlations between incremental and previous cumulative "small?"

The fifth criterion is to see whether there is an alternative explanation for a correlation between an incremental and the previous cumulative due to increases running down the accident years (which could be caused by increasing exposures or inflation). By fitting a simple linear accident trend, the chain ladder fails this criterion if for a substantive number of years the ratio is no longer a significant predictor.

Because of the small sample sizes, the assessment is not always obvious graphically, so we perform this test in the regression by looking at the ratio fitted last to see if it still has explanatory power.

The first two questions are assessed individually for pairs of adjacent developments (individual regressions), while the latter two questions are assessed globally (across all residuals). The final question is related to question 2 and again assessed on adjacent pairs of developments individually.

When assessment is performed on individual pairs of developments, only the first six pairs of developments are considered (there are too few points after that for a reliable assessment).

1) *Zero intercept:* Least squares line on graph of incremental vs. previous cumulative has intercept that is plausibly near origin; 68% of triangles failed on many years (3-6), (all triangles failed at least some years).

2) *Incremental significantly correlated with previous cumulative*: 52% of triangles fail on many years (3-6), 68% fail on several years.

3) *No CY trend changes*: 64% of triangles display strong CY trends, 28% display moderate CY trends, and 8% display weak indication of (plausibly random) CY trends.

4) *No trend in residual vs. fitted*: 32% show strong trends vs. fitted (linear or quadratic), 28% show moderate trends vs. fitted. 40% show little trend vs. fitted values.

Overall, 96% of triangles fail substantively (strong residual trends or fail on many years for the first pair of criteria) on at least one of these four criteria. Four percent show at least partial success for the chain ladder assumptions.

5) *Relationship not just a proxy for accident trends:* 84% substantively fail (that is, the relationship between y and x is substantively accounted for by a simple increasing or decreasing trend in the accident period direction, which indicates that exposures or inflation should be brought into the model). A further 4% partially fail (relationship remains for a few periods). In 4% the relationship for several years is strongly quadratic (a failure of the chain ladder assumption of linearity). In 8% there is remaining linear correlation between incremental and previous cumulative after accounting for a simple AY trend in incrementals.

Some of these criteria are plainly subjective (such as those based on appearance of diagnostic plots), but those effects classified as "strong" have deficiencies that are quite plain (though there's no clear dividing line between strong and moderate effects, so some people may well classify some of the moderate effects as strong or vice-versa); there is also room in the moderate/weak borderline for disagreement. Nevertheless, even by extremely generous criteria, still only a fraction (far less than a quarter) of the triangles considered could be even remotely regarded as suitable for the chain ladder.

If the results of this small study were to hold more generally, it seems that relatively few triangles will satisfy the basic assumptions about the mean for ratio models. Note that there are actually further assumptions required, such as variance assumptions – for the chain ladder, the assumption is that the variance is proportional to the mean. We have seen in other examples that these assumptions are not always tenable, so the proportion of triangles for which the model assumptions are plausible may be substantially lower than the present study might suggest.

It seems the only prudent course is to assume that ratio models do not describe a triangle, unless it has been clearly established that the model is a plausible description of the data.

SECTION 4: PROBABILISTIC TREND FAMILY MODELS

The problems that are regularly identified in the ratio models lead us to try to solve them.

i) We need to be able to deal with calendar-year trends (such as superimposed inflation) *and* to be able deal properly with the effect that multiplicative effects such as inflation have on the variance. This is one of many aspects that lead us to move from considering merely a log-link (as in the quasi-Poisson model) to a model with variance proportional to the square of the mean (of course, variance assumptions should be checked for reasonableness, as with any other assumption).

ii) The intercept term in the ELRF models is often necessary, but once calendar- and accident-year trends have been modeled, even in a very crude way, ratios rarely contribute anything further. This suggests that we should model the development year levels (corresponding to intercept terms) as well as effects in the calendar- and accident-years' directions.

iii) The overparameterization issue suggests that we should take advantage of the smoothness typically seen in the development period direction by relating the levels. Similarly, the periods of constant inflation or stable accident year levels tend to be fairly long-lived and we should be able to take advantage of that fact. (Further, in the full version of the model it is possible to relate/smooth accident years in a more extensive way than presented here.)

4.1 The basic model

All of these issues together point us toward what we call the Probabilistic Trend Family of models (*PTF*), which has the suggested features. It models incremental payments on the log scale. This allows inflation to be modeled appropriately (inflation impacts current and future payments, not past payments; consequently, cumulative payments and incurred are unsuited to models dealing with inflation). These may be normalized for a measure of exposure and adjusted for inflation before taking logs. The model described here may be fitted in a regression package (even in Excel), but predictive distributions of aggregates is somewhat involved.

The PTF approach models data as four components:

$$\text{data} = \text{development trend} + \text{accident trend} + \text{calendar trend} + \text{random}. \tag{4.1}$$

The first three components model the mean and the final component models the variability about the mean.

(i) *development trend*: The structure for modeling the runoff in the development year direction is as follows:

$$\delta_j = \gamma_1 + \gamma_2 + \ldots + \gamma_j \ . \tag{4.2}$$

The gamma parameters (γ_j) represent the shift in mean between the previous development and the current one. This allows models to adapt to shifts in level.

Setting gamma parameters to be equal allows the efficient modeling of constant trends. In a particular case, the fitted development year trend (runoff), $\hat{\delta}_p$, looks like this:

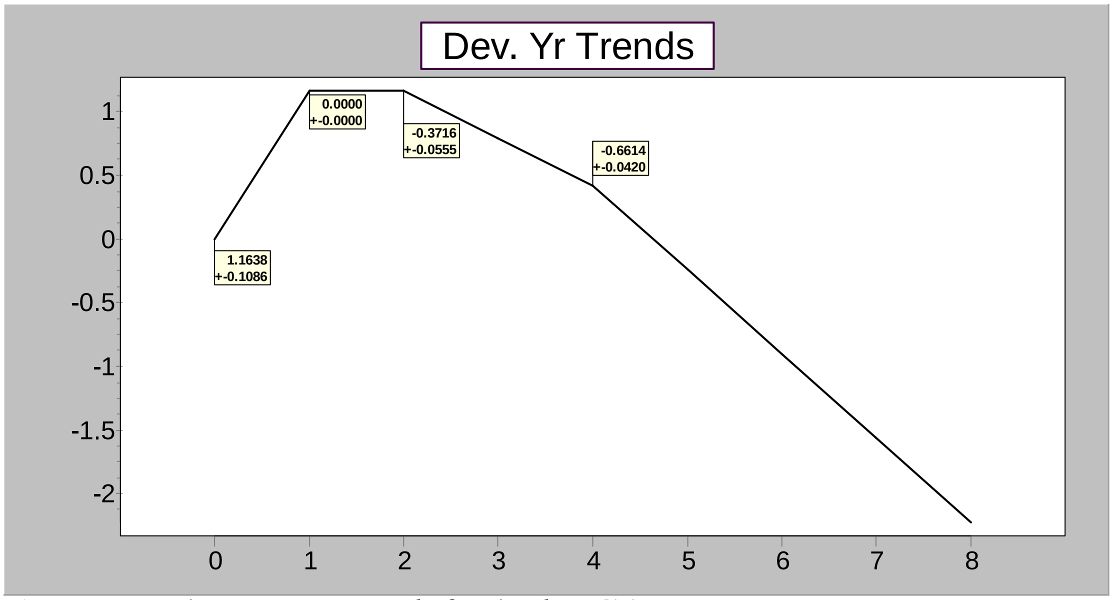


Figure 1: Development year trends for the data CTP

The level in development year zero is represented by a parameter alpha (α), and the gamma parameters represent (percentage) shifts to each development period after that.

For a single accident year, the average log(payment) in development j is:

$$E(y_j) = \alpha + \gamma_1 + \gamma_2 + \ldots + \gamma_j \ . \qquad (4.3)$$

In the case where there are no fitted trends in the mean in the other directions (accident year, calendar year), the model for the mean for the entire array can be of the same form:

$$E(y_{ij}) = \alpha + \gamma_1 + \gamma_2 + \ldots + \gamma_j \ . \qquad (4.4)$$

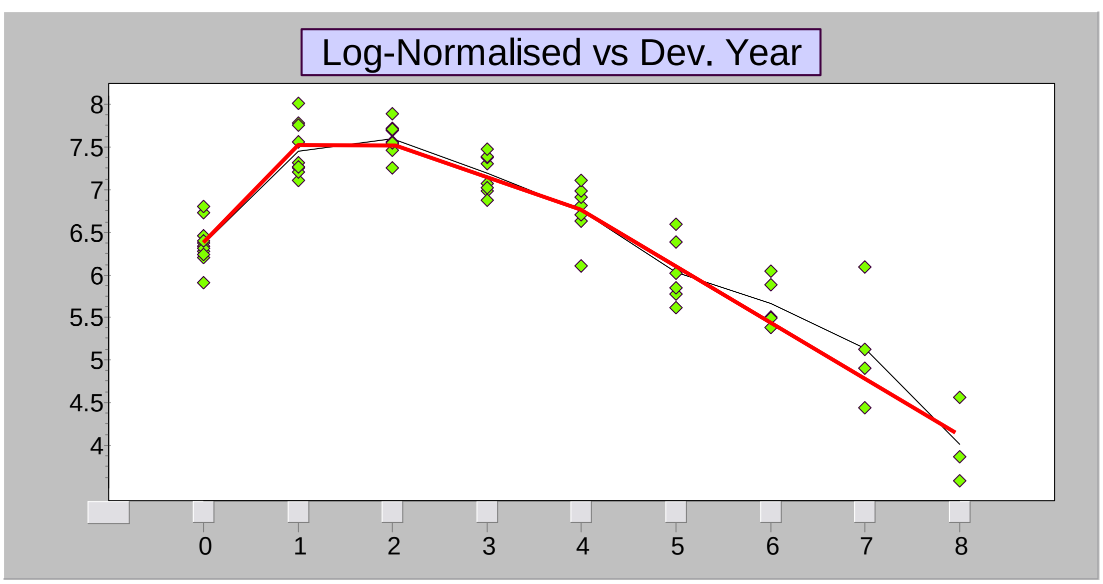


Figure 2: CTP data vs. development year, showing the fitted trend of model type (4.4)

(ii) *Accident trend*: Changing accident-year trends are possible by allowing the level for the accident year to change. The level of development year 0 in accident year *i* is α_i. After fitting calendar trends, the accident period trends are often fairly stable for long periods of time. Consequently, consecutive α- (or level-) *parameters* may be set to be equal.

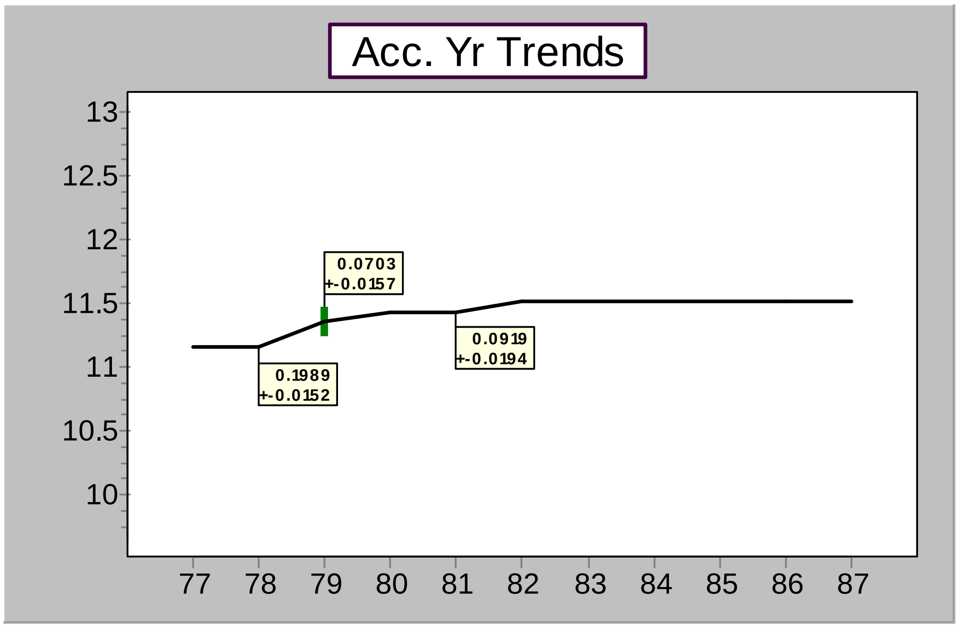


Figure 3: Accident year trends for the triangle ABC. Flat line segments show where consecutive parameters are equal.

Note that we can't show the fit to the data in this direction (because of effects in the other direction – development year effects are almost always present), unless we remove the effects of the other directions from the data (partial residuals).

Further, it sometimes makes sense to allow non-adjacent accident years to have the same level. For example, it is desirable when there has been a temporary shift in the level of paid losses that then returns to the original level.

(iii) *Calendar trend*: In real triangles we frequently find periods of stable, constant, or near-constant inflation, with sometimes abrupt changes. This can be modeled in similar fashion to the development period trend:

$$\kappa_t = \iota_2 + \iota_3 + \ldots + \iota_t \ . \tag{4.5}$$

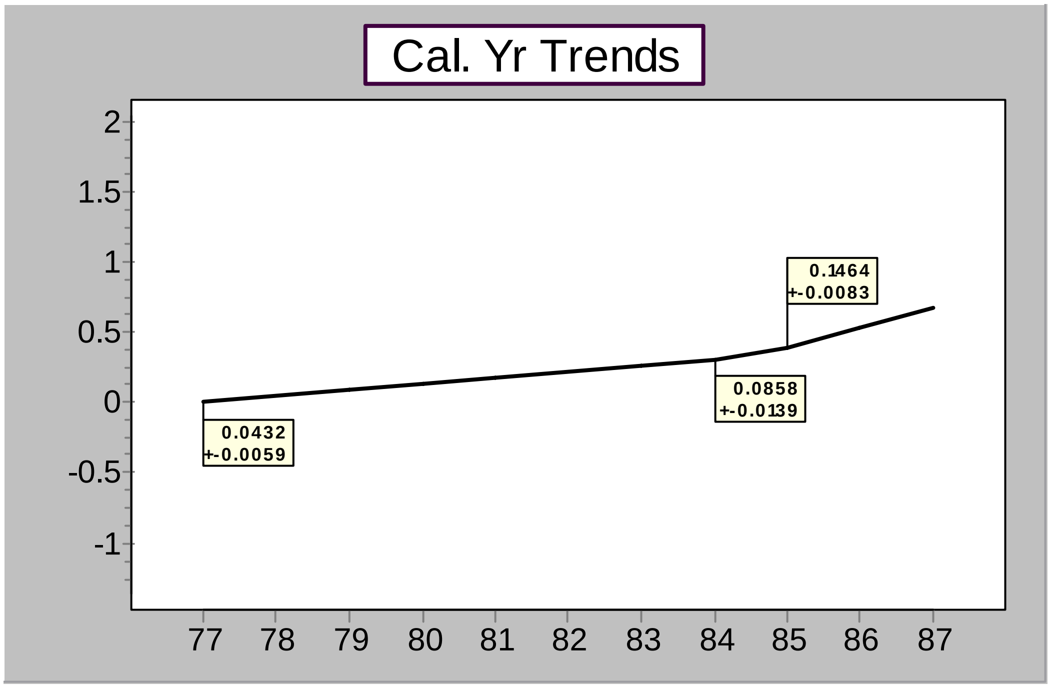


Figure 4: Calendar-year trends for the triangle ABC

Multicollinearity and the projection of trends to adjacent directions

Note that while it might sometimes be of interest to also have linear trends in the accident year direction (corresponding to periods constant growth in the accident year direction), it's not possible to simultaneously estimate trends in all years in all three directions at once, due to multicollinearity.

This is because of a fundamental property of triangles – trends in any of the directions project onto the adjacent direction(s). Consequently, it is necessary to restrict the potential to have a linear trend in one of the directions – in our case, the accident-year direction. A single linear trend in the accident year direction may be picked up by a trend in the calendar-year direction.

(iv) Random component:

The random component aims to describe everything of interest about the data not described by the mean. The distribution of the data about the mean (also called the distribution of the *error term*) is an essential part of the model. The errors are assumed to be independent, and by default the variance is taken to be constant on the log-scale (constant coefficient of variation).

Combining the components

This leaves the basic form of the Probabilistic Trend Family of models as:

$$y_{ij} = \alpha_i + \delta_j + \kappa_{i+j} + \varepsilon_{ij} \;;\quad \varepsilon_{ij} \sim (0, \sigma^2),$$

where each of the terms describing the mean has a particular structure in order to allow for the tendency for the trends or levels to be stable. This equation is of the form *data = accident trend + development trend + calendar trend + random*, as with equation 4.1. Each component is parameterized so as to be able to provide parsimonious descriptions of the trends typically seen in real data.

An assumed error distribution is not required in order to estimate parameters, but estimation is via least squares, which is optimal in the case of normality. In order to calculate explicit predictive distributions, we make the assumptions that the ε_{ij} are normal (that is, that the original p_{ij} are lognormal), but this assumption should be checked (since if it's implausible, the predictive distributions will not have the desired coverage probabilities). So we explicitly assume that ε_{ij} are normal, and so the original data are assumed to be lognormal.

We summarize this model in terms of four pictures, representing the four components. This gives an instant visual understanding of what is going on in the model. When combined with residual plots, the practitioner is able to rapidly assess both the components of the model for the mean and the characteristics of the data about the mean. For example, here is a reasonable model for a particular set of data (ABC, analyzed in detail in Barnett and Zehnwirth, 2000):

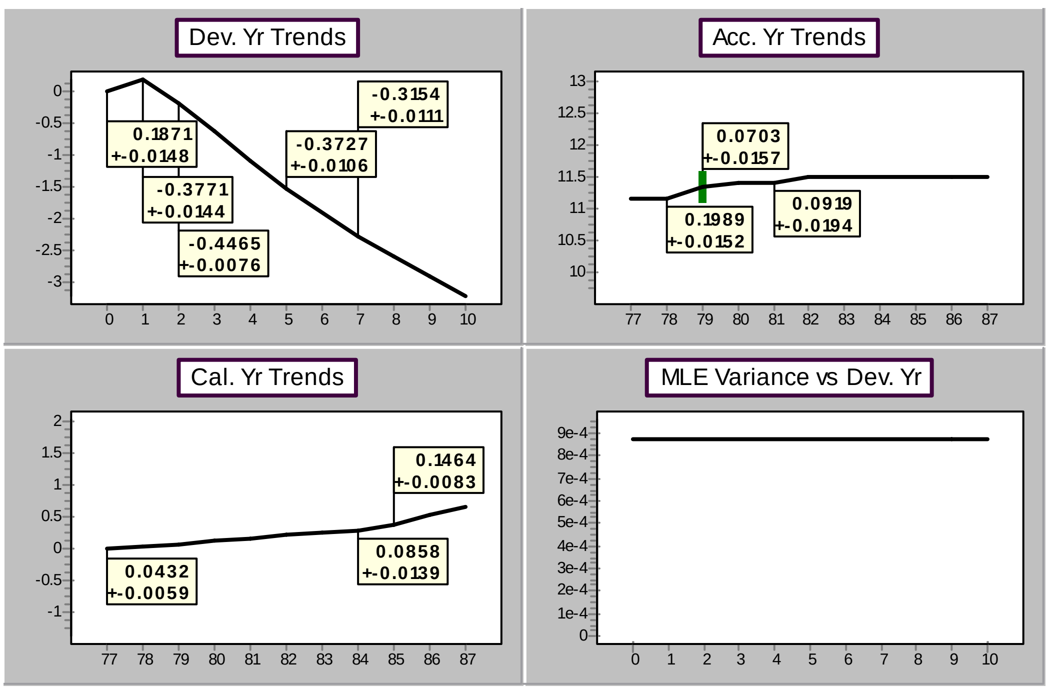


Figure 5a: Model trends in the three directions and fitted variance.

The residuals from this model are below.

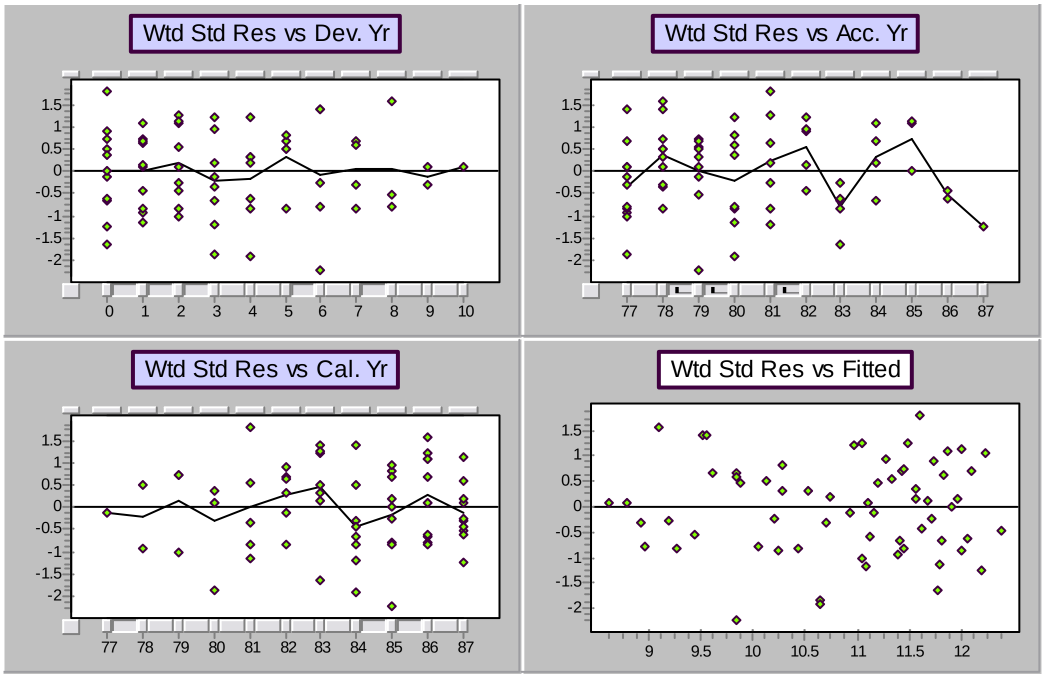


Figure 5b: Residuals vs. the three directions and against fitted values.

There is some pattern remaining in the accident-year direction that is better captured with a model that smoothes the accident-year trends rather than one that adds further discrete parameters; the modest remaining movements don't support separate parameters.

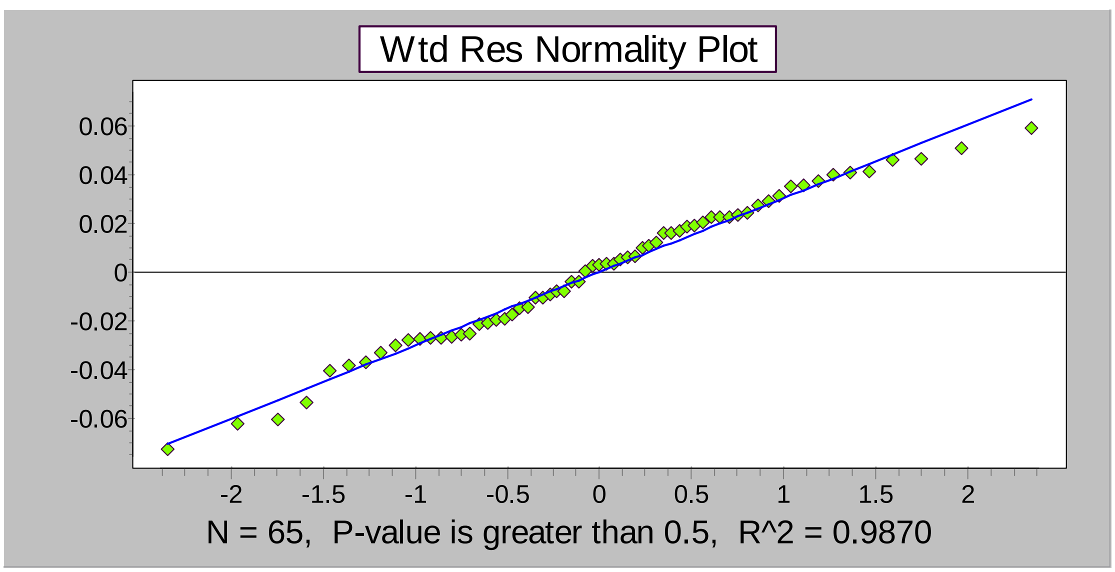


Figure 6: Assessment of the normality of the residuals. The lack of curvature indicates that the data are quite consistent with the assumption of normality.

4.2 Does this model family fit the data?

Being designed to describe features we regularly find in loss triangles, the model family includes many models that describe some triangles well – but each triangle is different, so a particular model should be chosen to describe the data at hand. Model critique via diagnostic displays and statistics should always be carefully assessed.

Indeed, the rich PTF model family includes in it a model whose linear predictor is identical in form to that of the quasi-Poisson GLM. Note that both the PTF models and the quasi-Poisson GLM are based on describing the mean on the log-scale, and that a PTF model with $\kappa_t = 0$ for all calendar years and with all α_i and γ_j different has an identical description of the mean (that is, as a two-way cross classification structure in logs) as a quasi-Poisson GLM. Further, as noted in section 2.2, the Mack-type chain ladder model also (necessarily) has parameters in the same places (though in that case, the accident-year parameters are implicit, being contained in the conditioning on the first development). Consequently, if there is a circumstance in which the PTF models fail to describe the mean, then a chain ladder-like model *is also guaranteed to fail*. On the other hand, it is frequently the case that a PTF model will succeed in describing the data where a ratio model fails.

Numerous examples are given, for example, in Barnett and Zehnwirth (2000), where triangles that are not well described by a chain ladder or other ratio model are well captured by a PTF model.

4.3 Extending the Basic Probabilistic Trend Family

Modeling changes in variance:

While the constant (log-scale) variance assumption is often tenable for many triangles, sometimes the variance is not constant. However, when it occurs, while the two are related, often the variance is seen to change more clearly with the development period than with the mean. When it is not reasonably constant, the variance is often relatively stable for a number of years, and so when necessary we allow the variance to change with development, while allowing adjacent variance parameters to be equal.

As mentioned before, a more complex form of smoothing (a generalization of exponential smoothing) is sometimes used (notably in the accident-year direction) in the PTF framework, but we will not pursue details of that since it requires more specialized algorithms than the simple regression methods required for the model described here to implement.

There are a variety of other ways in which this model can be extended. For example, the same structure on the mean could be used on count data in a GLM framework.

In the remainder of the paper we describe an extension of the PTF framework to the analysis of multiple triangles (such as different types of claim – asbestos vs. non-asbestos, different territories, several excess of loss layers, or different lines of business.

SECTION 5: MULTIVARIATE PTF MODELS

We relate PTF models by allowing for two different type of relationship between triangles:

(i) Payments may be correlated about their overall trends (if one is higher than average, the other may be higher than average).

(ii) The trends, or the differences in trends, may be related across triangles. For example, it is often the case with two related triangles that not only does superimposed inflation change in the same places in both, but the size of the change is similar in both triangles.

Where the individual models have the same *design* (the same pattern of observations with identical parameter structure), the combined model for both triangles taken together is a *generalized least squares* model (GLS). In this case, the parameter estimates are not affected by the correlations between triangles, and the calculations may be performed in two stages (estimation in individual triangles followed by estimation of the between triangle correlation).

While the correlations between triangles don't affect the estimates in GLS, they do affect the distribution of the forecast of the aggregate (or, indeed, the difference or even other functions if they occur).

More generally, the designs differ somewhat, and the combined model is a seemingly unrelated regression model (SUR). However, as noted in section 4.3, the models can actually get slightly more complicated than SUR.

As with the discussion of PTF models, we will concentrate on an SUR approach in this paper. Packages are available that can estimate SUR models, for example, the free statistical package *R* (http://cran.r-project.org/) has an extension package called *systemfit* which can fit SUR models.

SECTION 6: EXAMPLES

6.1 Relationships between two lines of business (LOB1/LOB3)

We consider a pair of triangles from two lines of business, which we label here as LOB1 and LOB3. They are full 10x10 triangles over the same period of time.

Here is a simple model fit for LOB1 without setting any non-consecutive parameters to be equal (though the pair of peaks against accident years suggests that it might be reasonable to pool information in that way).

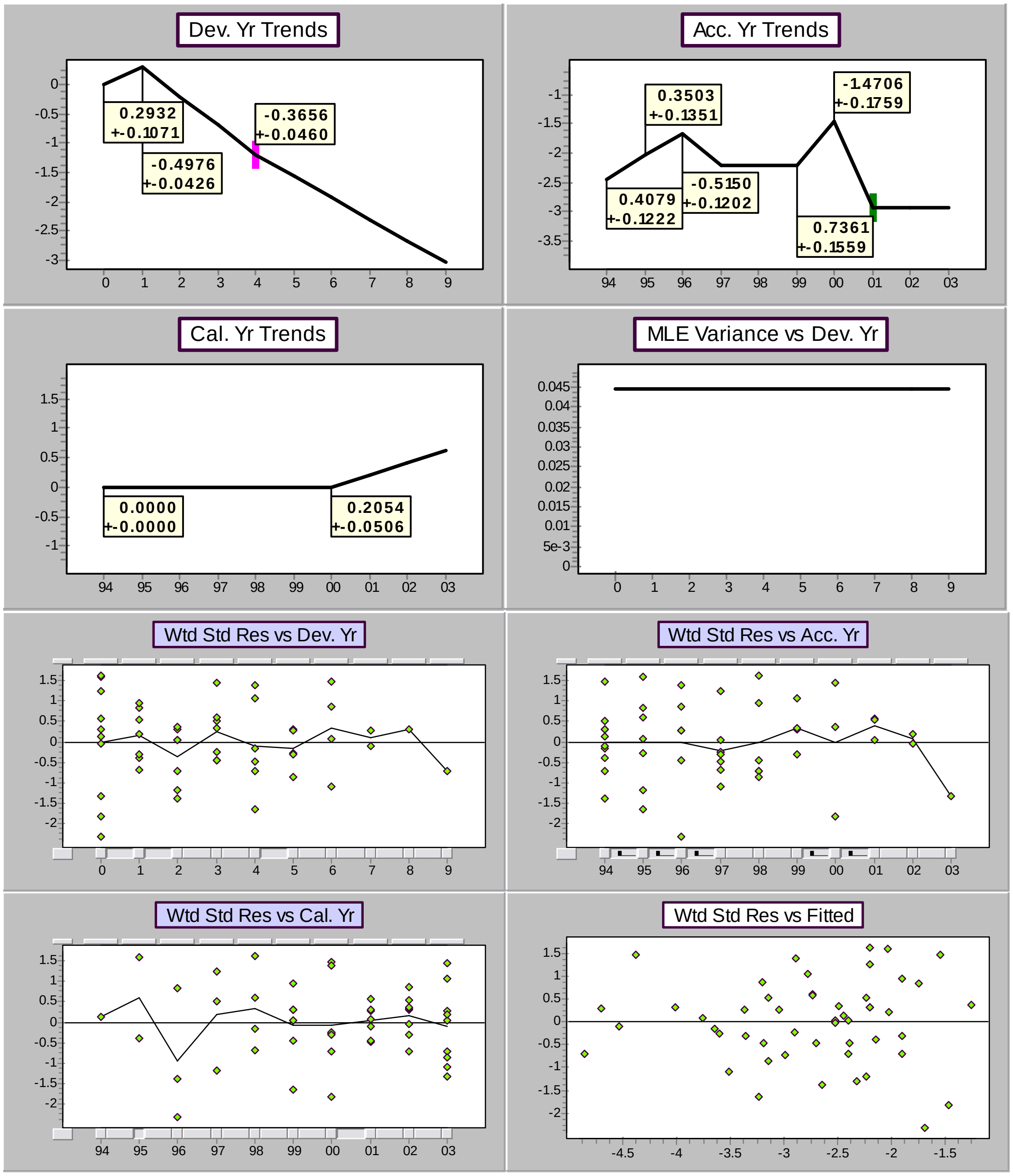


Figure 7: Model display and residuals for LOB1

We see this model provides a reasonable description of the data; it suggests that beginning in 2000, there was a dramatic (0.2054) increase in calendar-period trend (about 23% p.a.). While you should note that there was a substantial drop in the accident-year level in 2000, both the increase in calendar-year trend and the drop against accident year appear to be required for a reasonable description of the data.

It turns out that LOB3 has a major change in calendar- and accident-year trends that also occur in the same places as LOB1 (calendar-year 2000 and accident-years 2000 and 2001 respectively).

However, it is quite difficult to model LOB3 well, on its own. If we fit a reasonable model to the other directions, the accident year residuals appear as follows:

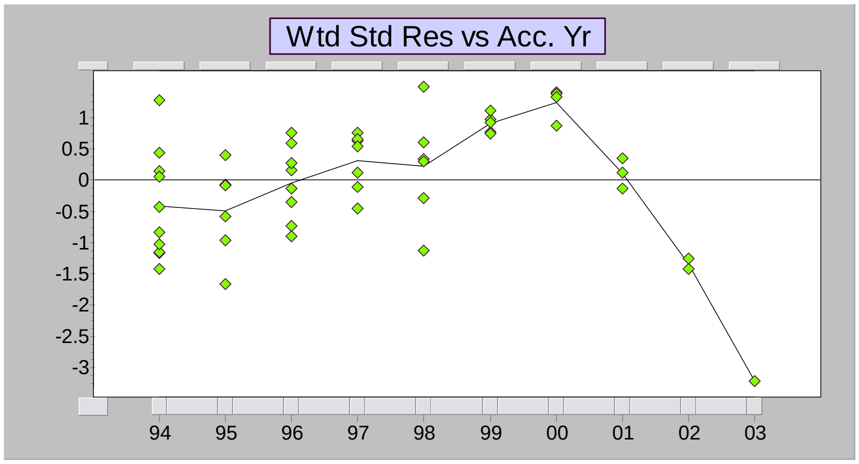


Figure 8: remaining trends in the accident year direction after removing trends in the other directions

The above residual display indicates a need to capture two linear trends against accident years. There are a variety of ways to model that with this data. One approach is to allow the accident-year levels to increase until the change, as with the following model:

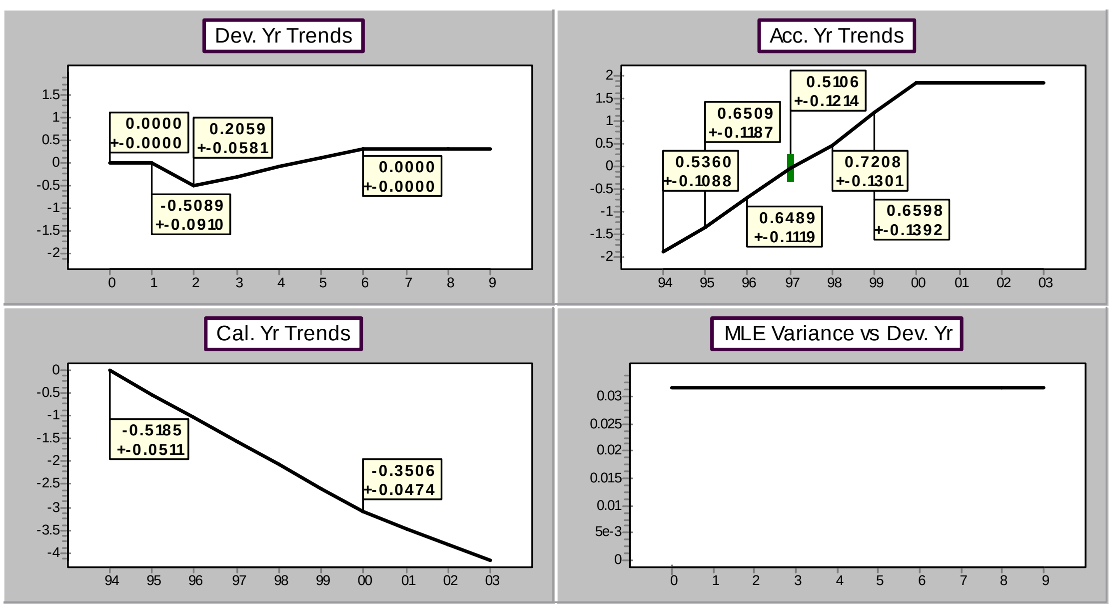


Figure 9a: An overparameterized model for LOB3

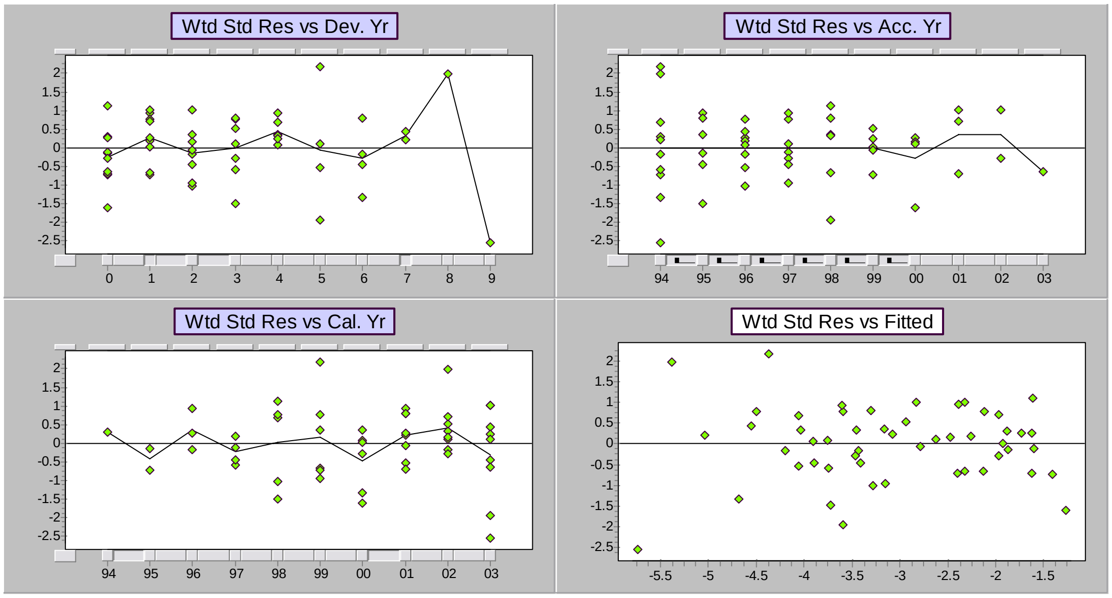


Figure 9b: Residuals from the above model.

This particular formulation of the model is somewhat overparameterized – though not nearly as much as the chain ladder, and it yields a substantially better description of the data. For example, see the residuals from a chain ladder fit in the calendar-year direction:

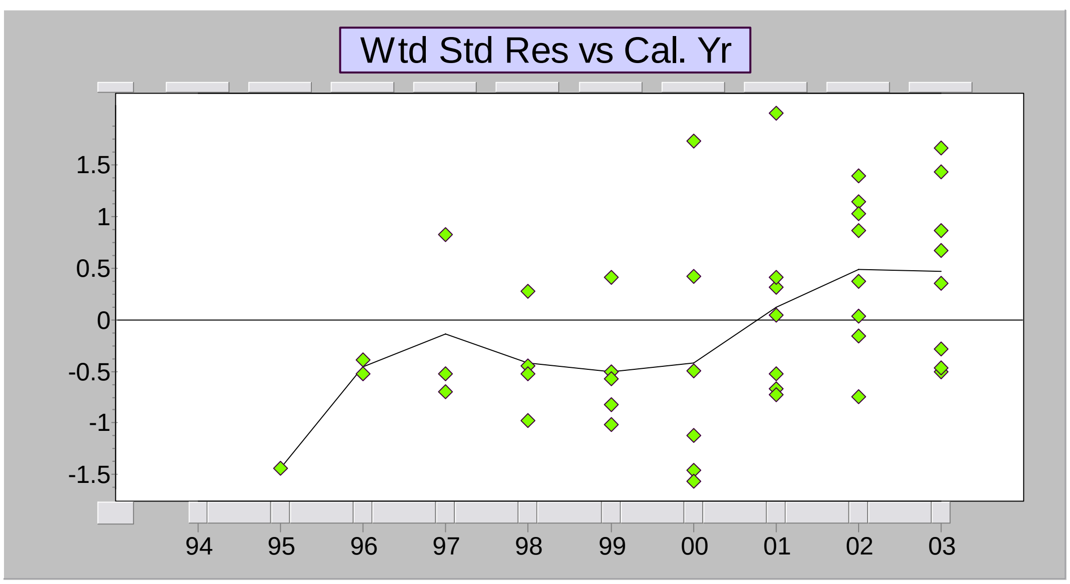


Figure 10: LOB3 Chain ladder residuals vs. calendar years

The graph shows distinct remaining calendar-year trends. Additionally, intercepts are necessary for several years (that is, ratios are inadequate). LOB1 shows similar uncaptured calendar-period trends when fitting ratios.

There is a certain amount of difficulty in obtaining simple descriptions of the combined effect of the accident- and calendar-period effects in LOB3. However, when the triangles are modeled together, it becomes more straightforward. Let us model both triangles together:

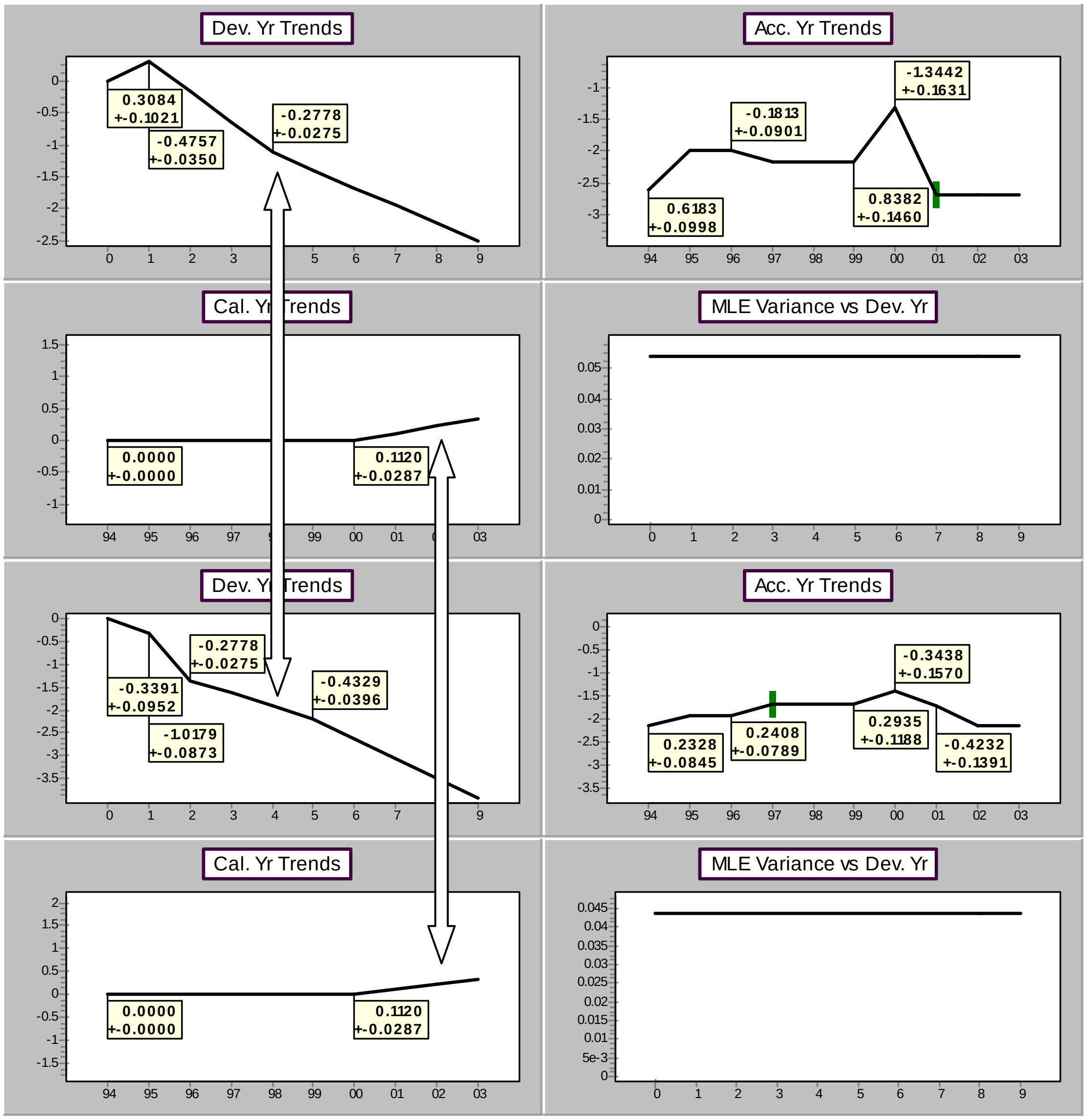


Figure 11: Composite Model for LOB1(top) and LOB3 (bottom). The arrows indicate equal trends.

In the model above, two development period trends (for overlapping periods of developments) and two concurrent calendar-period trends have been set to be equal. The accident-year level in the first and last accident years within LOB3 are also equal to each other.

The fit for each of the two lines of business has borrowed strength from the other..

This model has stabilized the estimates of trend very well, which has been particularly helpful for LOB3 – in fact (even when double-counting the common parameters), the number of parameters used to describe each array is smaller, yet the residuals are reasonable:

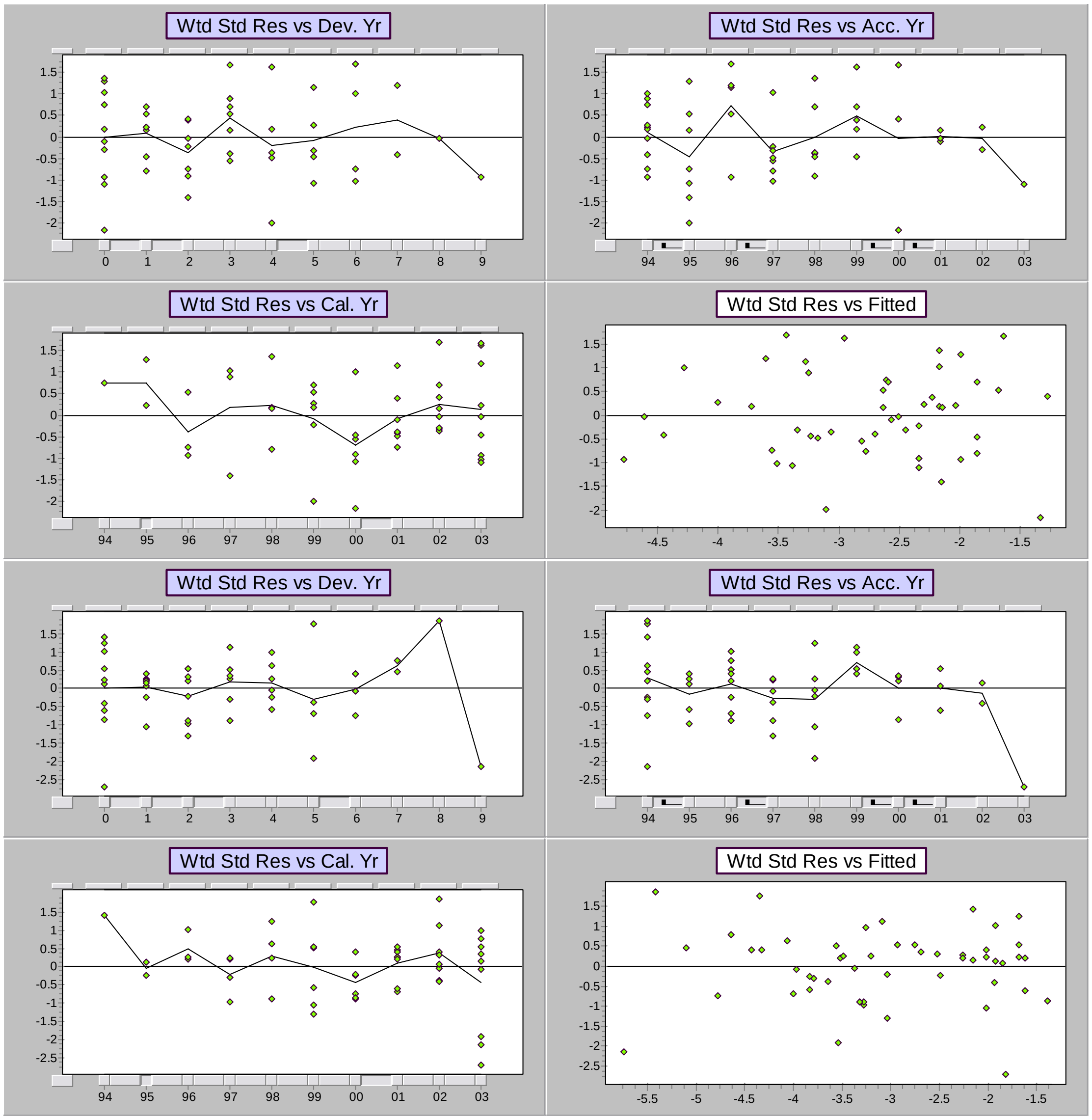


Figure 12: residuals for the composite model for LOB1 (top) and LOB3 (bottom)

Somewhat better results can be obtained if the smoothing of accident year trend estimates is used, but the mathematics is beyond the scope of this paper.

The estimated residual correlation between triangles is 0.39. This correlation can be seen in the moderate tendency for corresponding plots

Forecasts:

Note that the forecasts will be correlated in two ways:

(i) Parameter correlation: both the fact that two parameters will be shared (one development and one calendar period) between triangles and the fact the remaining parameters will be correlated across triangles contribute to the parameter correlation;

(ii) The model assumes that the data are correlated around the model. The estimate of the size of this correlation is roughly 0.4.

Let's examine a table of aggregate forecasts for the two lines and their sum.

Composite Model Forecast: Accident-Year Summary: 1 Unit = $1,000

	Mean Reserve	Ultimate	Standard Dev.	CV of Reserve
LOB1	705,717	2,362,497	95,613	0.14
LOB3	1,182,458	4,560,860	165,913	0.14
Aggreg.	1,888,175	6,923,357	244,212	0.13

The CV of the aggregate is smaller than the CV of the individual lines, as we would anticipate.

By way of comparison, the figures for the individual models have slightly higher coefficients of variation; while the aggregate of the means for the models fitted individually is almost identical (around half a percent smaller), the standard deviations are substantially reduced because of the better use of information (reducing the contribution of parameter uncertainty to risk capital).

We can compute forecasts and standard deviations of aggregates, but in order to compute quantiles (such as for value at risk calculations or some risk capital calculations) it is necessary to simulate.

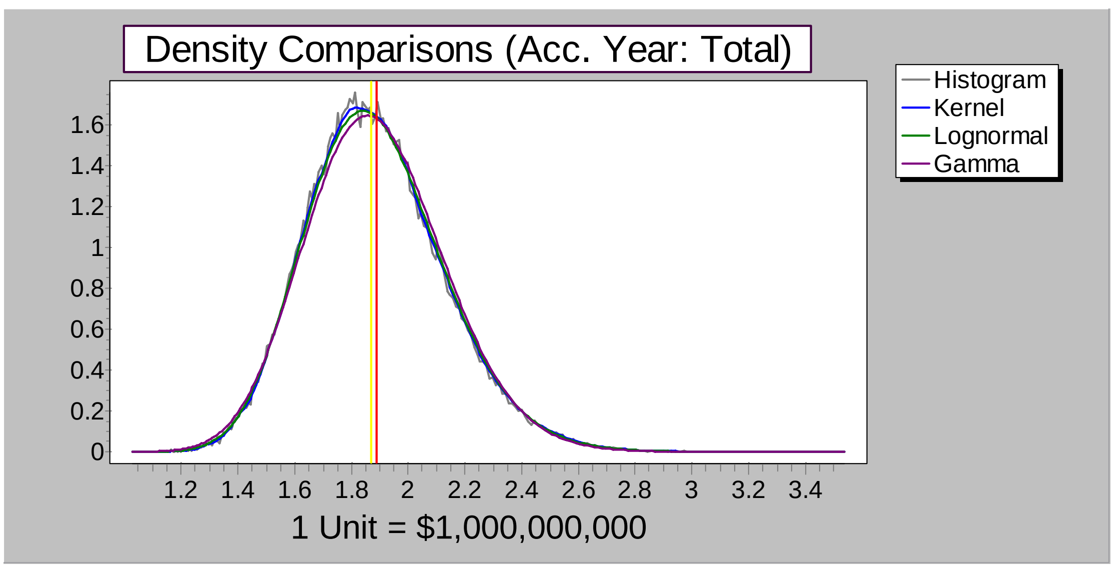


Figure 13: simulated distribution of aggregate of LOB1 and LOB3 based on 100,000 simulations. Rough grey line = simulated values, blue line is a kernel density estimate (smoothed simulated values), green curve = lognormal with the calculated mean and standard deviation, and purple = equivalent gamma density.

The PALD simulations enable the computation of quantiles, either directly from the simulated values or from one of the smooth curves.

LOB 1 LOB 3:Composite Reserve PALD Summary
Selected Quantile Statistics and Value at Risk (Acc. Year: Total) Unit = $1 Billion

%	Sample			Kernel			LogNormal			Gamma		
	Quantile	# S.D.s	VaR	Quantile	# S.D.s	VaR	Quantile	# S.D.s	VaR	Quantile	# S.D.s	VaR
90.0	2.205	1.298	0.317	2.208	1.308	0.319	2.209	1.312	0.321	2.207	1.306	0.319
95.0	2.313	1.738	**0.424**	2.316	1.751	0.428	2.315	1.746	0.426	2.307	1.715	0.419
98.0	2.443	2.271	0.555	2.446	2.284	0.558	2.440	2.259	0.552	2.423	2.190	0.535
99.0	2.532	2.638	0.644	2.536	2.653	0.648	2.527	2.616	0.639	2.502	2.515	0.614
99.5	2.616	2.981	0.728	2.619	2.993	0.731	2.609	2.953	0.721	2.576	2.818	0.688
99.9	2.796	3.717	0.908	2.805	3.755	0.917	2.788	3.685	0.900	2.734	3.462	0.845

Let's compare those results with corresponding results where the process correlations are set to zero:

LOB 1 LOB 3:Composite Reserve PALD Summary
Selected Quantile Statistics and Value at Risk (Acc. Year: Total) Unit = $1 Billion

%	Sample			Kernel			LogNormal			Gamma		
	Quantile	# S.D.s	VaR	Quantile	# S.D.s	VaR	Quantile	# S.D.s	VaR	Quantile	# S.D.s	VaR
90.0	2.246	1.311	0.292	2.248	1.322	0.294	2.246	1.310	0.291	2.244	1.303	0.290
95.0	2.342	1.744	**0.388**	2.345	1.756	0.391	2.340	1.735	0.386	2.334	1.707	0.380
98.0	2.457	2.262	0.503	2.460	2.274	0.506	2.451	2.235	0.497	2.438	2.174	0.483
99.0	2.537	2.618	0.582	2.541	2.639	0.587	2.528	2.581	0.574	2.509	2.492	0.554
99.5	2.611	2.951	0.656	2.617	2.978	0.662	2.601	2.907	0.647	2.575	2.789	0.620
99.9	2.793	3.772	0.839	2.796	3.783	0.841	2.757	3.610	0.803	2.714	3.417	0.760

Compare, for example, the VaR at the 95th percentile for the fitted model with the estimated correlation is $424 million, whereas if the two triangles are assumed to have zero process correlation, the VaR is $388 million. If process correlation is not taken into account, the VaR is $36 million (8.5%) too small. Similar effects are seen with TailVar, though the effect (in percentage terms) is a little larger at a given level (due to looking further into the tail). *Economic capital requirements are underestimated if the correlations between the predictions are not incorporated.*

6.2 Excess-of-loss type layers

In this example we consider the relationships between two layers net of reinsurance (losses limited to 1M and losses limited to 2M), and the layer corresponding to the difference between them, (1M excess of 1M). If the analysis was done at the scale of the original dollars, the three arrays would be linearly dependent (the first and third should add to the second), but on the log scale this is no longer the case.

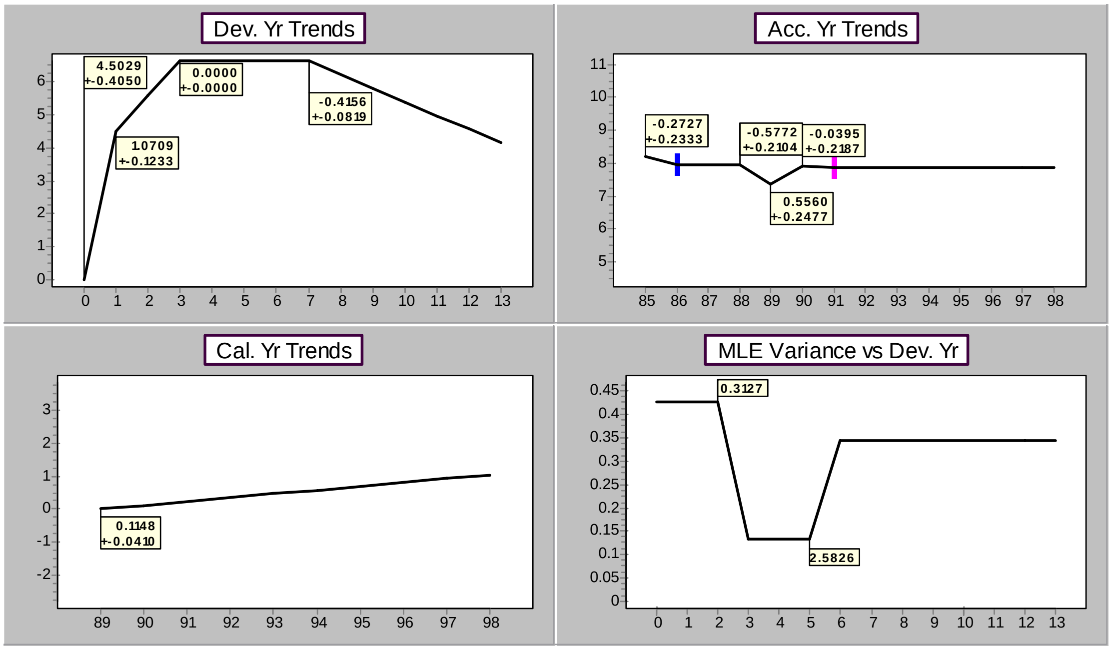


Figure 14a: model displays for layer 1M

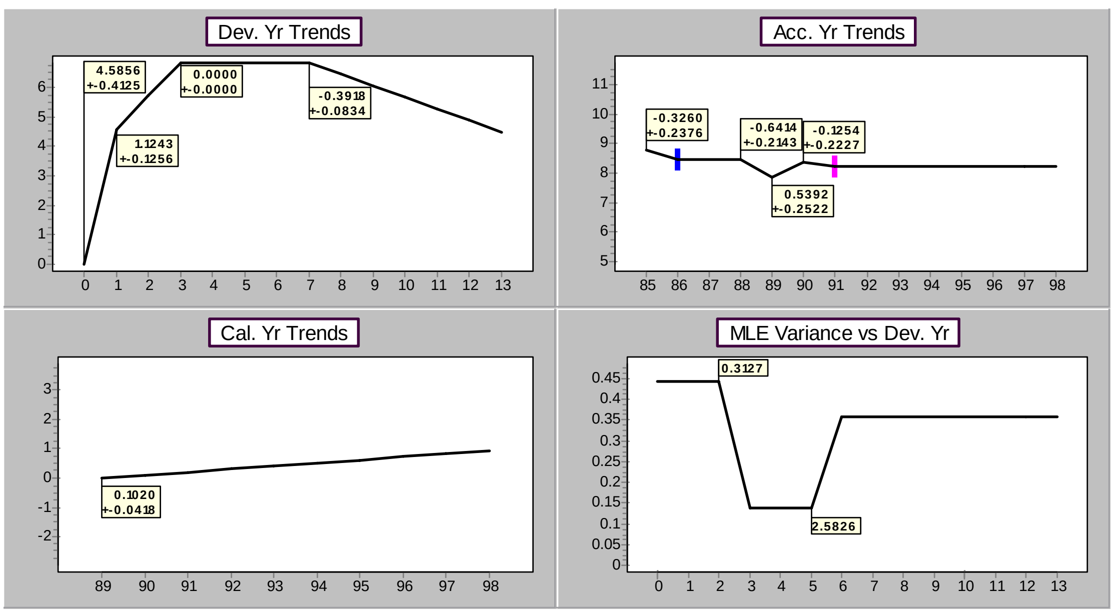


Figure 14b: model displays for layer 2M

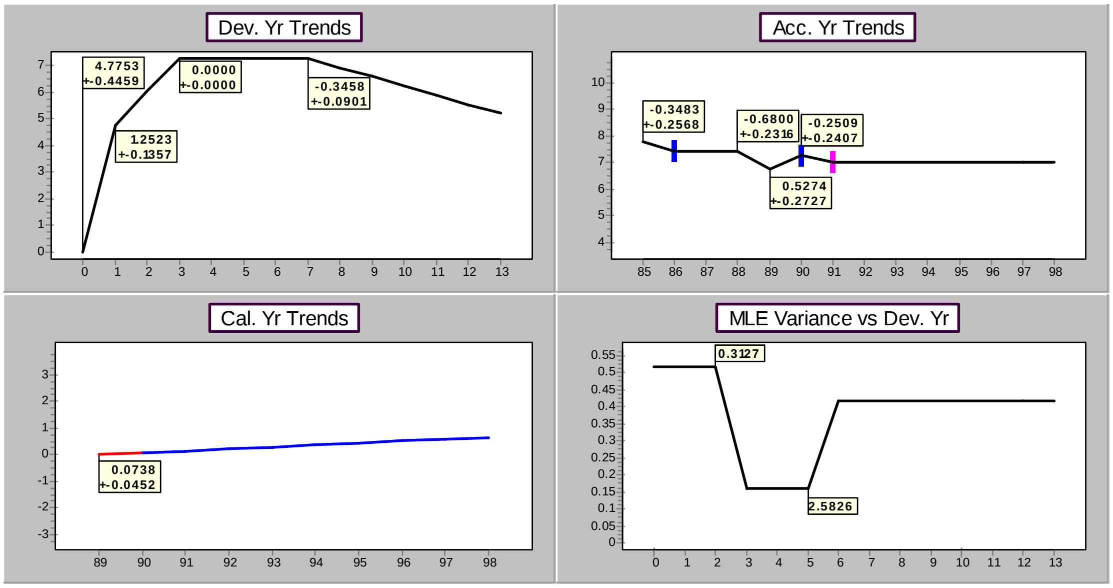


Figure 14c: model displays for layer 1M XS 1M

We see that the model chosen for the two ground-up layers and for the excess layer are quite similar, but the calendar-year trend for the excess layer (the third model plot) has a large standard error (there is high uncertainty about the underlying value) – indeed we can't be sure there's a non-zero trend there at all! On the other hand, if we're projecting that trend out into the future, we don't want to set it to zero; its estimate corresponds to an annual inflation of about 7.5% per annum.

It would be useful if we could borrow trend information from the ground up layers to help estimate some of the trends, in which case the estimate of the calendar-period trend may be better estimated.

As it turns out, a suitable combined multivariate model of the same form does exactly this. The *difference* in accident-year level from 1988 to 1989 is very close for the layer to 1M and for the next layer (and further, for the layer to 1M, the levels either side of the down-and-up bump are very close, though this has much less impact). If we borrow strength across the first two layers by setting that change in level to be equal, the standard error on the calendar-period trend comes down substantially, and the estimate is now extremely close to zero (0.0026), or about a quarter of a percent. With this model, it makes little difference to the forecasts whether or not the estimate of superimposed inflation for that layer is set to zero.

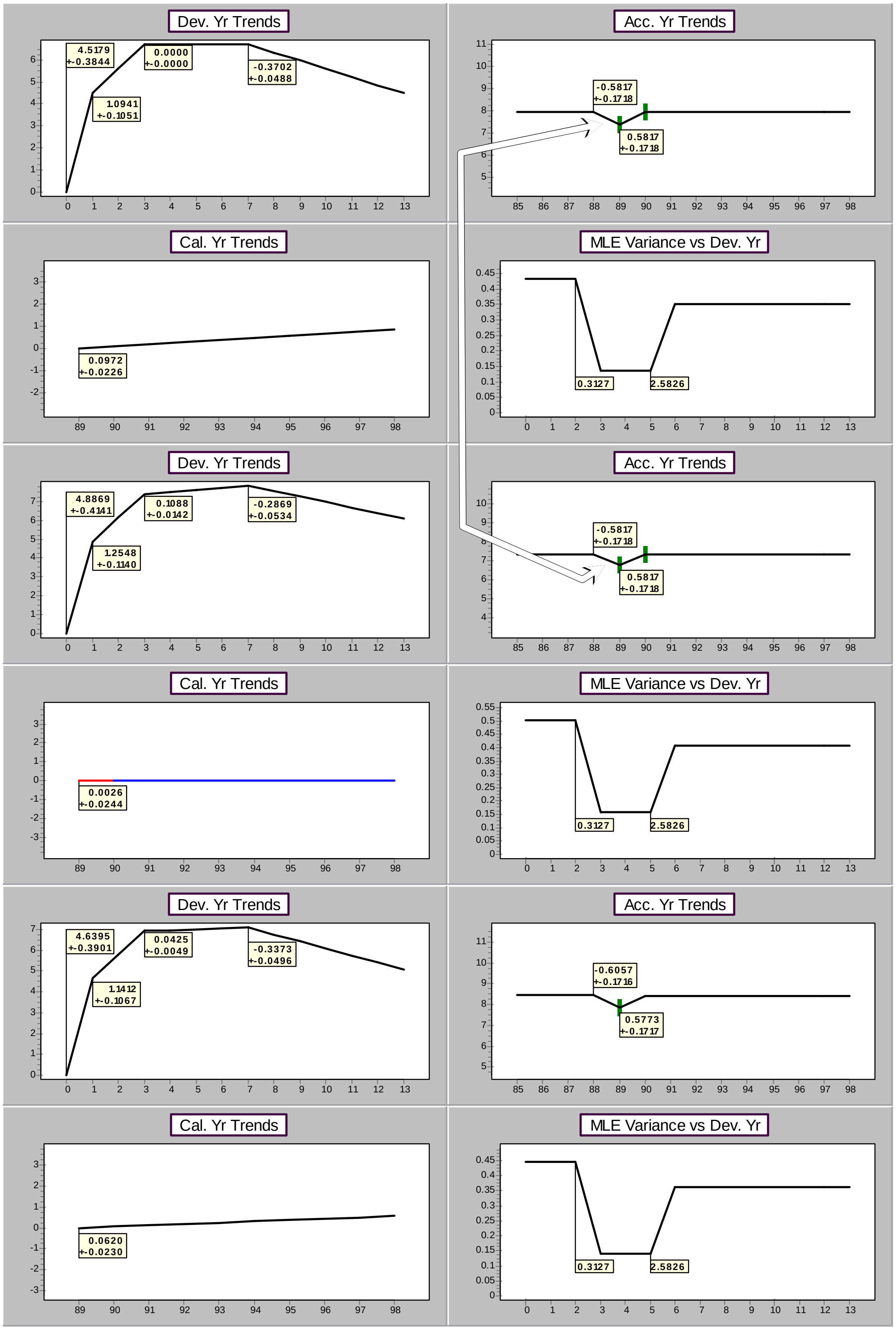
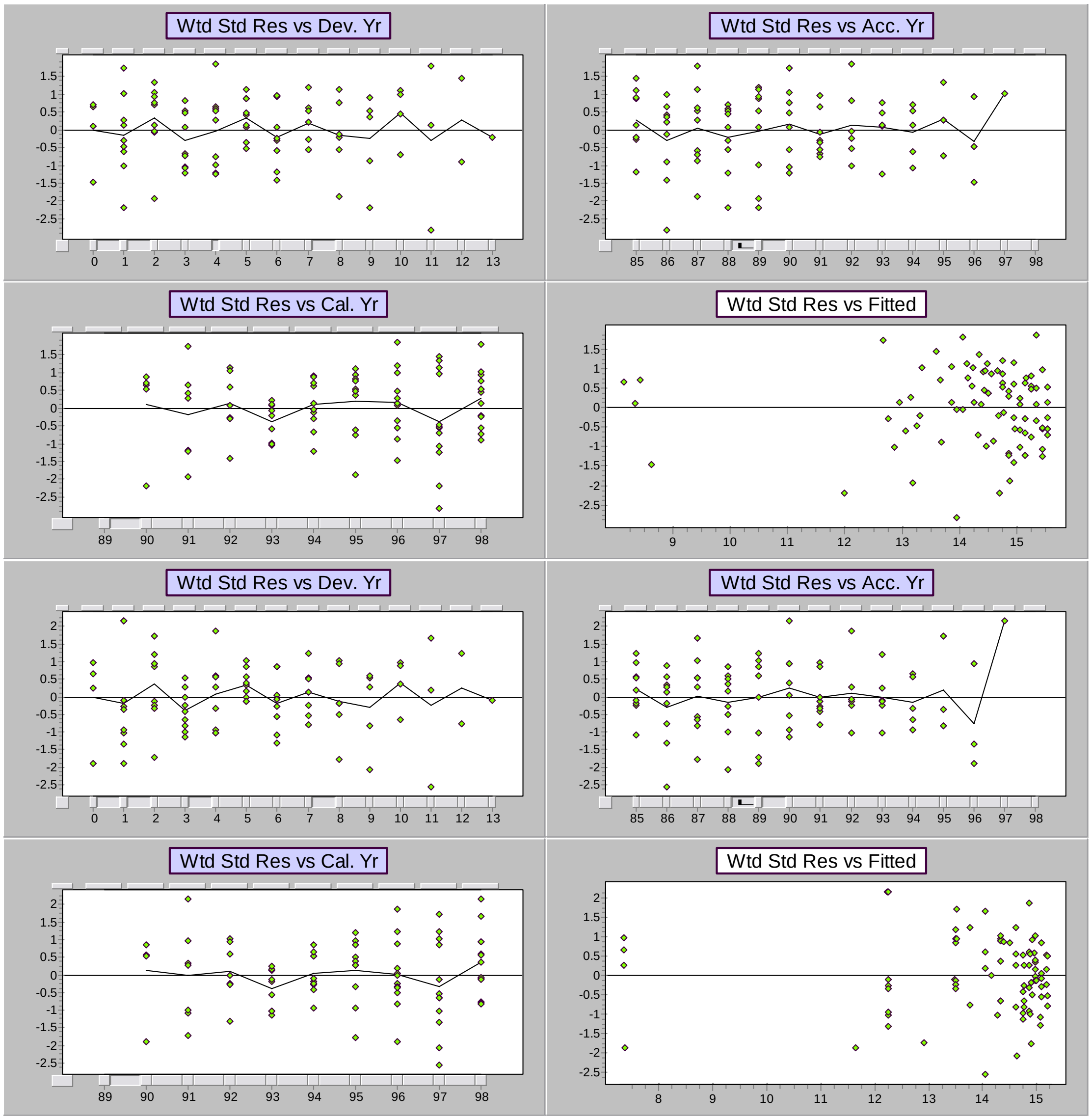


Figure 15: Model displays for combined model for layers 1M, 2M, and 1MXS1M. The arrow indicates a shift in accident-year level that is set equal across layers.

Here are the residuals from the above model



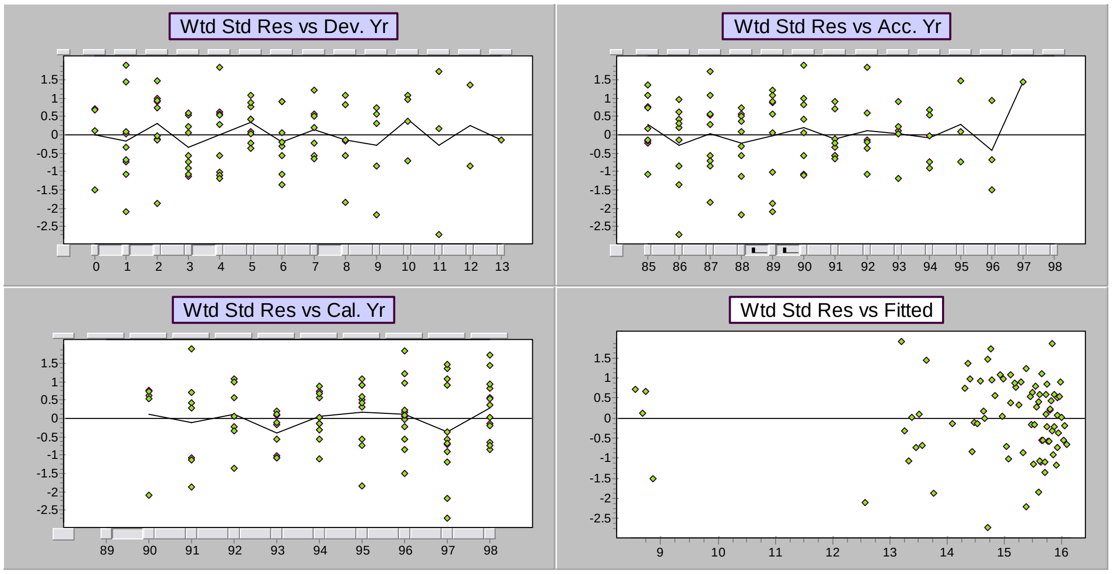


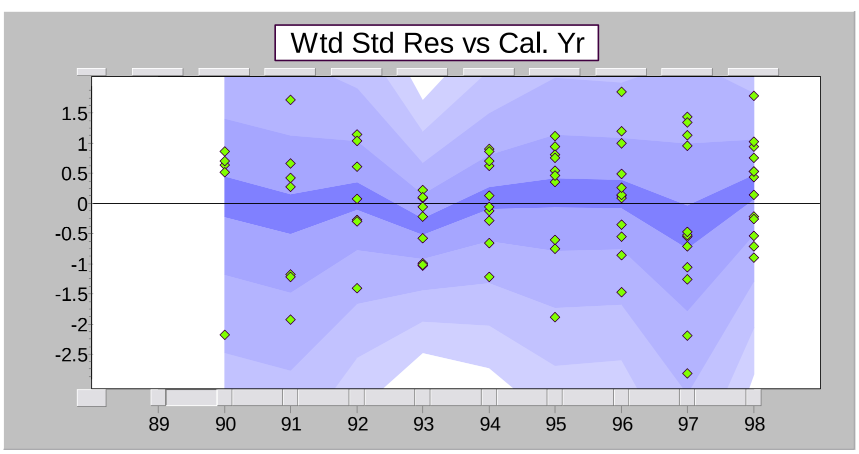
Figure 16: Residuals from the fitted composite model

The residuals indicate a reasonable fit for this model. The residuals do tend to "move together (perhaps not surprisingly!), and this is reflected in the correlations, which are very high.

Correlations between triangles:

	1M:PL(I)	1MXS1M	All 2M
1M:	1	0.960	0.995
1MXS1M:	0.960	1	0.982
2M:	0.995	0.982	1

These correlations can be seen by comparing corresponding residual displays across triangles. For example, in the residual displays below, the blue contours show multiples of standard deviations either side of the means by year.



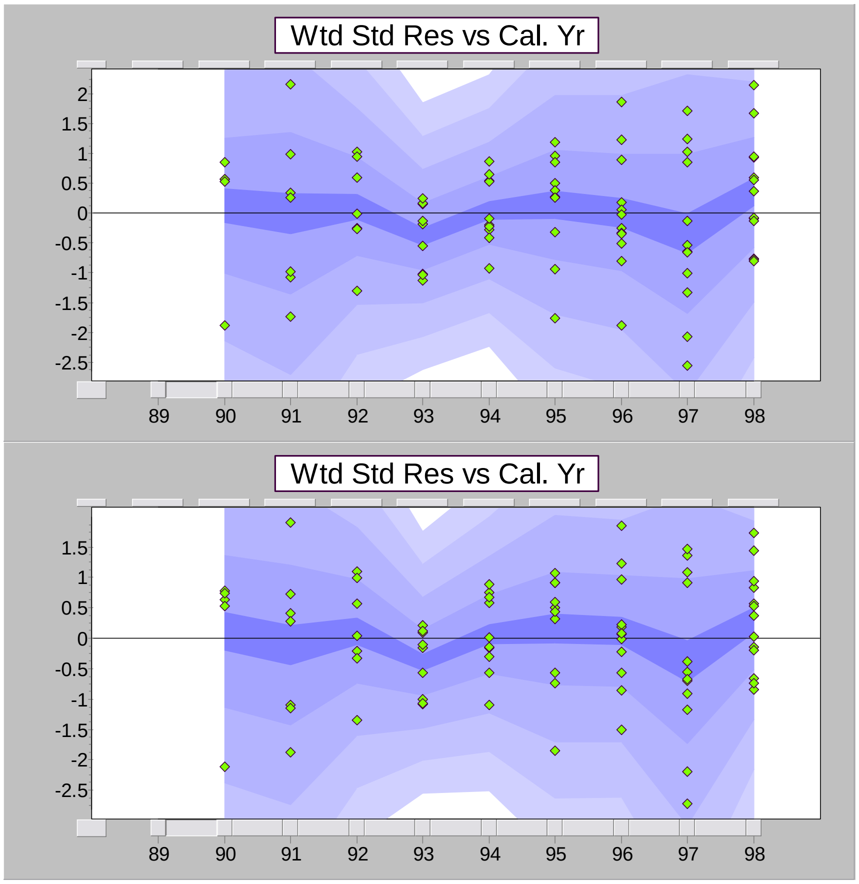


Figure 17: Residuals vs. calendar year for 1M, 1MXS1M, and 2M.

The "wiggles" in the blue lines are similar across each triangle, because of the high correlations between corresponding points. The tendency of the residuals to move up and down together is their correlation.

Here is a summary of the results from the forecasts with this model:

1M 1Mxs1M 2M:Composite Reserve Forecast Summaries:
Accident Year Summary 1 Unit = $1M

				CV	
	Mean Reserve	Ultimate	Standard Dev.	Reserve	Ultimate
1M	454	692	92	0.20	0.13
1MXS1M	244	421	47	0.21	0.11
2M	677	1,113	138	0.20	0.12
1M+1MXS1M	678	1,114	138	0.20	0.12

Note that adding 1M and 1MXS1M gives essentially the same answer as 2M, as we would wish. This consistency is preserved across a variety of reasonable models for this data.

The CV for the 2M and for the smaller layer 1M is almost the same.

Moving from retaining losses below 2M to losses below 1M (reinsuring 1M XS 1M) doesn't help the cedant's CV at all!

6.3 Gross vs. Net data

The following data is gross and net of Excess of loss-type reinsurance. Analysis of gross and net data together can help to assess the value of the reinsurance, and aid the design of reinsurance better suited to the direct insurer's needs.

Here is a display of a good fitted model for the gross data.

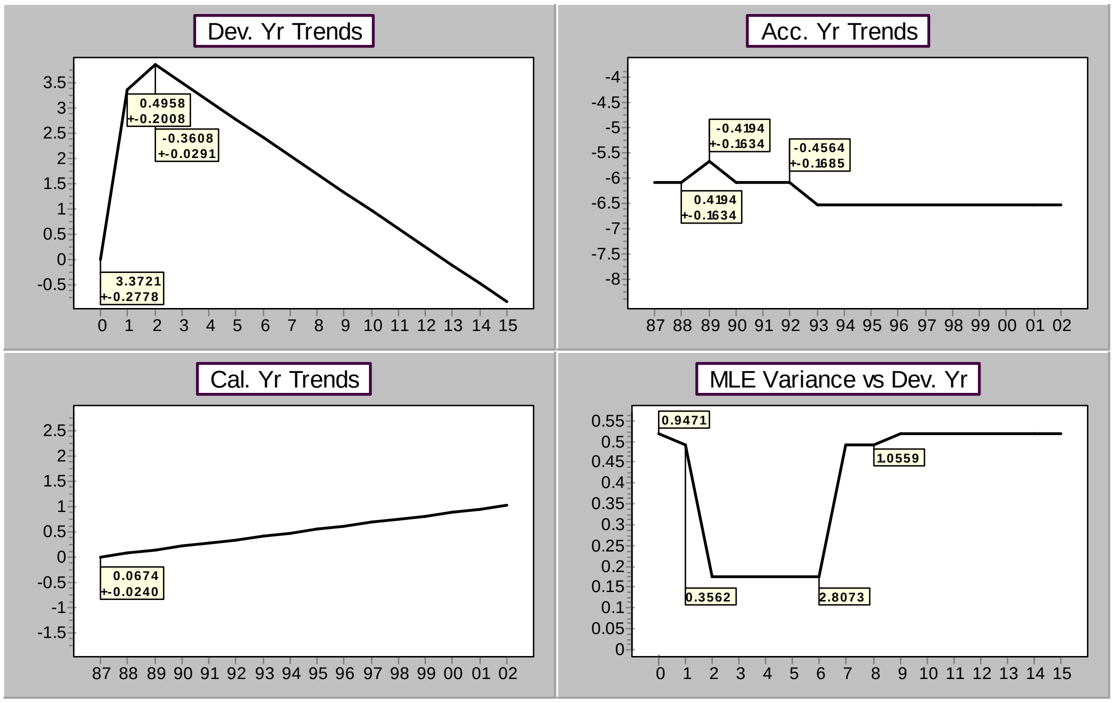


Figure 18: Model display for the Gross data.

There is smoothing of accident-period levels between 88 and 89 and between 92 and 93. The accident-year level for the 1988 (and earlier) mean is set equal to the 1990-1992 mean. Further note that the model has the level of variance of observations at developments 2-6 substantially lower than for the other years (heteroskedasticity). There's a stable trend of around 7% per annum ($e^{0.0674}-1$) in the calendar-year direction.

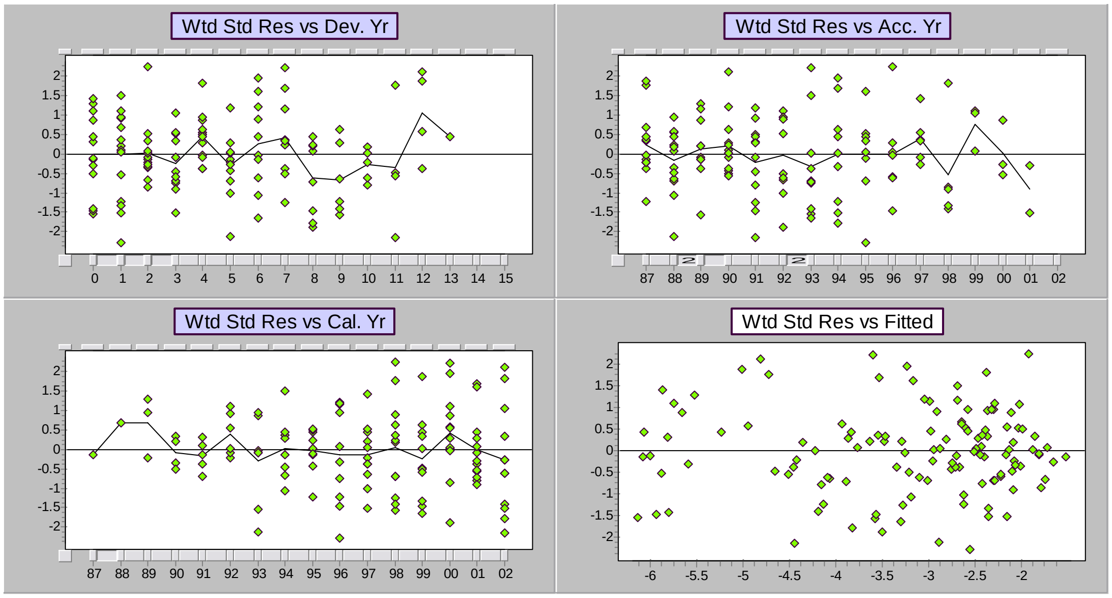


Figure 19: Residual plots for the model. The fit to the calendar- and accident-year directions is good; there is some lack of fit in the tail (where payments are very low).

Now let's model the Net data:

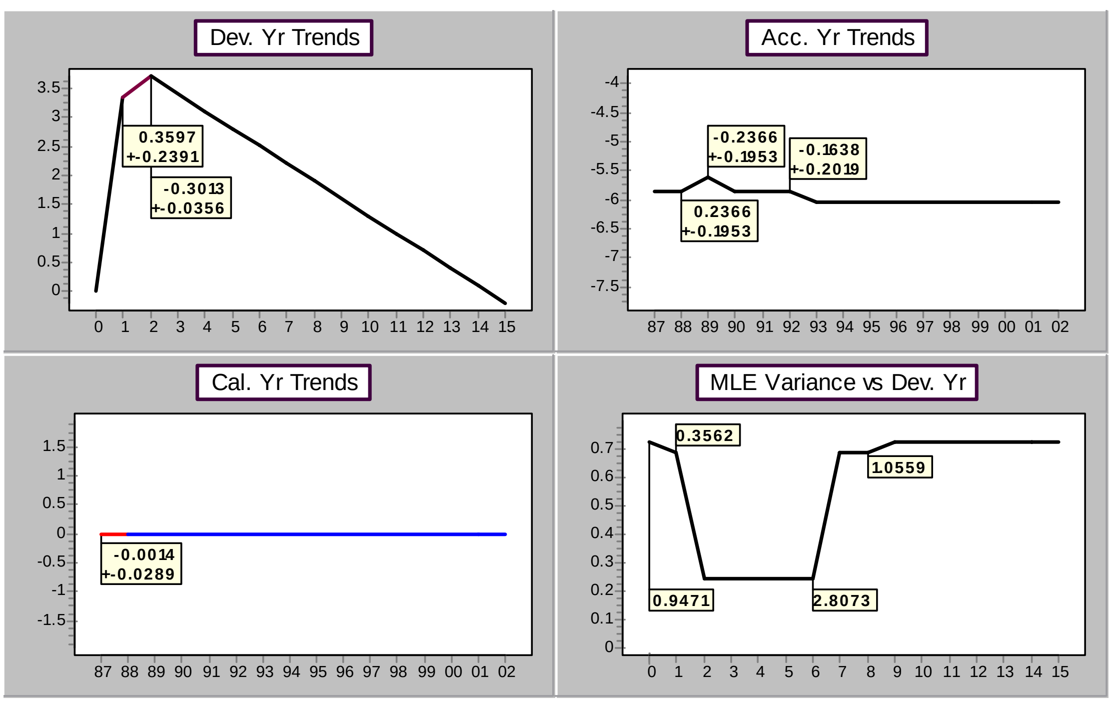


Figure 20a: Model display for the Net data

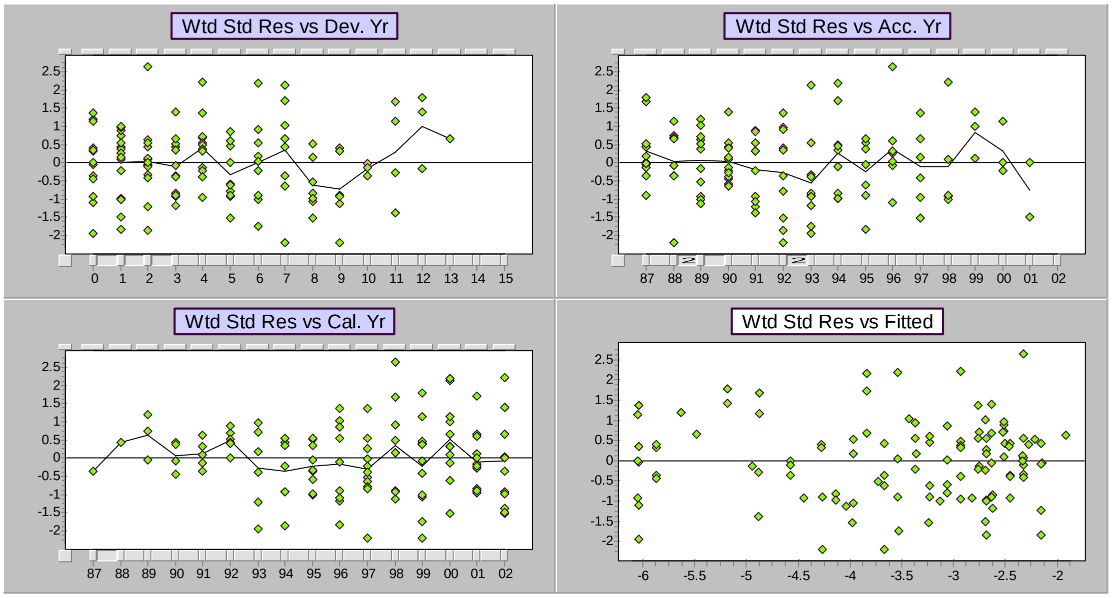


Figure 20b: Residuals for the Net data

Here, the model is quite similar to the gross, but as we see, the estimated calendar-year trend is essentially zero (0.14% with a standard error of about 3%). The fit is again, reasonably good.

The original insurer seems to be entirely ceding the growth (calendar-period inflation). While not particularly surprising, this is important information for both the cedant and the reinsurer.

In Figure 21, below, the graph shows the relationship between residuals for the two triangles. The correlation between the residuals of these models is quite high (this is common with reinsurance data), at around 0.81, and the normality assumption is reasonable.

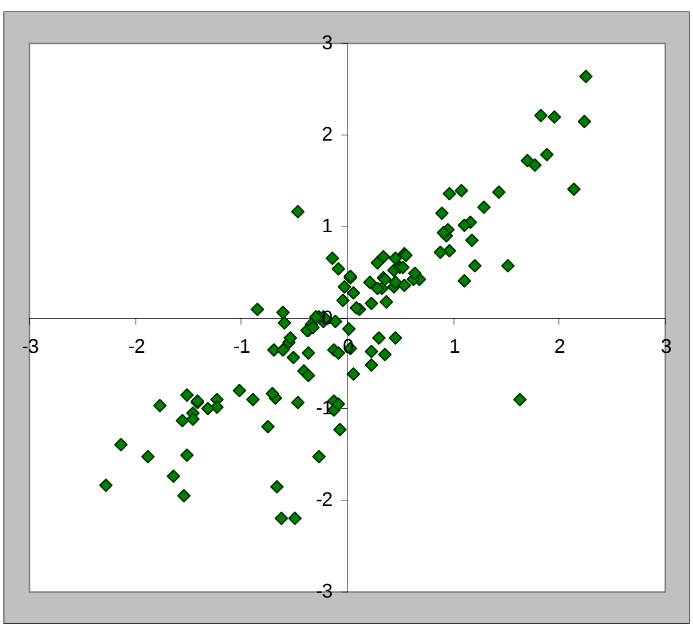


Figure 21*:* Plot of corresponding residuals from the two models, Net vs. Gross

Combined model:

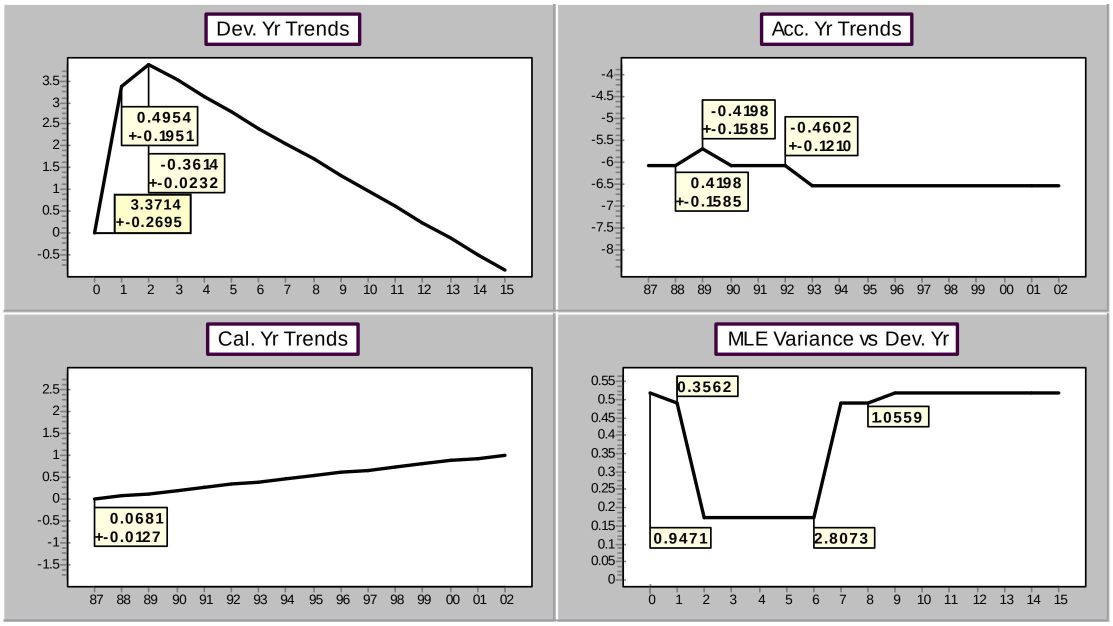


Figure 22: Model display for the combined model for the Gross data

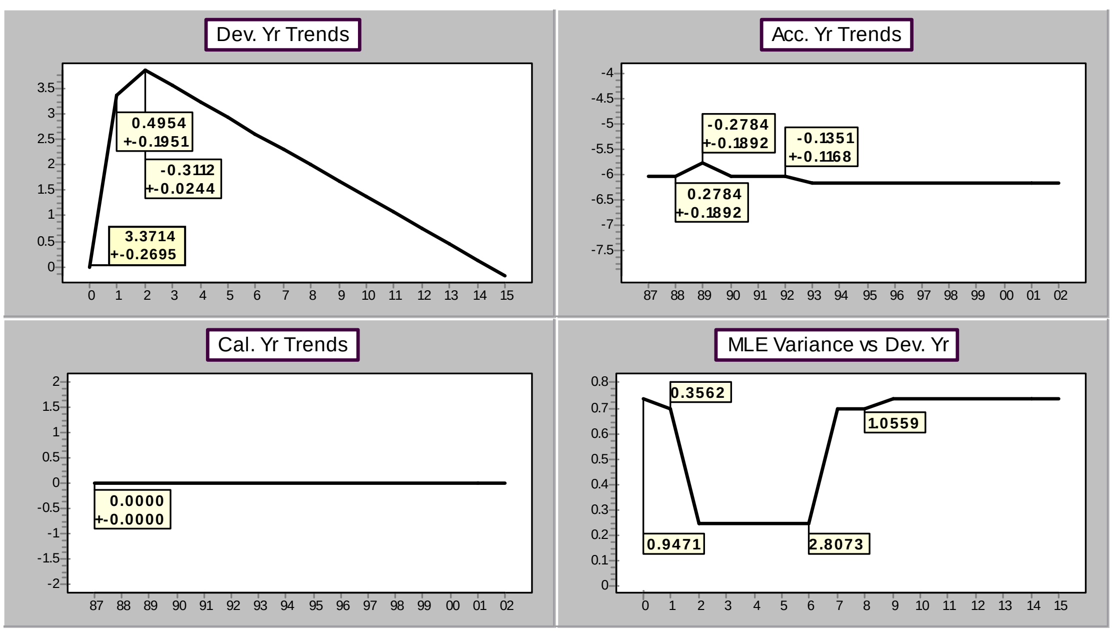


Figure 23: Model display of combined model for Net data

The first two development period trends had high uncertainties and have been set to be equal (they're effectively percentage changes in level across development periods).

Correlation in residuals is a little higher at around 83%, but otherwise the models are similar. The estimated trend for the calendar years for the Net data became even smaller (and less uncertain), and was set to zero (though it would make little difference if it was retained in this case, just as with the previous example).

Forecasting this combined model yields an interesting result:

Comp N & G: Composite: Accident Year Summary

				CV	
	Mean Reserve	Ultimate	Standard Dev.	Reserve	Ultimate
:FAC ENG G:	101,705	216,641	15,450	0.15	0.07
:FAC ENG N:	52,821	134,323	8,725	0.17	0.06

The CV for the Net data (0.17) is slightly higher than for the Gross (0.15)! Even taking into account the fact that some of the trends will be less certain for the Net data (which can pull up the CV), this seems to suggest that the reinsurance is not achieving the goal of reducing the riskiness for the cedant.

DISCUSSION

These studies have some very important consequences for capital risk charges for both claims liability and underwriting. The forecast standard errors considered here include process variability, which is a major component of the risk, and which we cannot reduce below its inherent level; we can, however, reduce parameter uncertainty.

The fact that (percentage) changes in trend are often closely related across related triangles means that we can estimate critical components, such as calendar-year trends with less uncertainty – and so predictive distributions can become more concentrated. At the same time, we must consider the impact of the relationships between the triangles – relationships between parameters (including equal trends or trend changes) mean that the forecasts are correlated, and then there is the contribution of the correlation of the error term across triangles. These correlations contribute to the estimate of risk and so are critical to the estimation of economic capital.

Between some lines of business, after incorporating common trends, there's often little residual correlation, but as we saw in section 6.1, we can certainly get substantial correlations – neither close to independent nor close to very high dependence—and this can make a substantive difference to required capital. Consequently, using either the assumption of independence or the assumption that the dependence attains its upper bound could lead to either a radical under- or over- assessment of the required risk capital.

REFERENCES

[1] Ashe, F., "An essay at measuring the variance of estimates of outstanding claim payments," *ASTIN Bulletin* 1986, 16(S):S99-S113.

[2] Barnett, G. and B. Zehnwirth, "Best Estimates for Reserves," *Proceedings of the Casualty Actuarial Society* 2000, 87(167):245-321.

[3] Barnett, G., B. Zehnwirth, and E. Dubossarsky, "When Can Accident Years Be Regarded As Development Years?" *Proceedings of the Casualty Actuarial Society* 2005, 92(177):239-256.

[4] Barnett, G., B. Zehnwirth and D. Odell, "Meaningful Intervals," Paper submitted to the CAS 2008 Reserves Call Paper Program, Casualty Actuarial Society *E-Forum*, Fall 2008, http://www.casact.org/pubs/forum/08fforum/.

[5] England, P. D. and R. J., Verrall, "Predictive Distributions of Outstanding Liabilities In General Insurance," *Annals of Actuarial Science* 2006, 1(2):221-270.

[6] Hachemeister, C. A. and J. N. Stanard, "IBNR Claims Count Estimation with Static Lag Functions," Paper presented at the 1975 Spring Meeting of the Casualty Actuarial Society.

[7] Mack, T., "Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates," *ASTIN Bulletin* 1993, 23(2):213-225.

[8] Murphy, D., "Unbiased Loss Development Factors," *Proceedings of the Casualty Actuarial Society* 1994, 81(154):154–222.

[9] Renshaw, A. E. and R. J. Verrall, "A Stochastic Model Underlying the Chain-Ladder Technique," *British Actuarial Journal* 1998, 4(4):903-923.