A Model to Test for and Accommodate Reserving Cycles

Thomas S. Wright, MA, CStat, FIA

Abstract

In recent years several commentators have noted evidence for a "reserving cycle" linked to the underwriting cycle. It seems that in many classes of non-life insurance, when premium rates are relatively low, claim development patterns tend to be longer-tailed than when premium rates are high. If this is the case, then traditional reserving methods based on an assumption that the development pattern is the same for all origin years will tend to overstate reserves for periods where premium rates were high, and understate reserves for periods where premium rates were low. The present paper reviews the evidence for a reserving cycle and discusses possible causes. A mathematical model is then proposed that accommodates the main possible causes. The purpose of this model is three-fold: (a) to test for the existence of reserving cycle effects, (b) to help identify the causes, and (c) to produce improved reserve estimates. An example analysis is presented using the proposed model. The evidence for the existence of reserving cycles is now sufficiently strong that, in the author's opinion, it is important for reserving actuaries to be aware of the possibility of cyclical effects, to investigate evidence for such effects in any reserving exercise, and (where there is strong evidence) to adjust reserve estimates accordingly. The model proposed in the present paper can be implemented in Excel and will often be a useful tool for these purposes.

Keywords. Reserving cycle, underwriting cycle, development patterns, curve fitting, least squares, premium rate indices.

1. INTRODUCTION

1.1 Research Context

1.1.1 Bob Conger's presentation at GIRO 2002

The idea of a "reserving cycle" was first given prominence by Bob Conger (then CAS President) in his keynote presentation to the 2002 GIRO Convention in the UK. For all classes of US property/casualty insurance combined, and for workers compensation alone, he showed the ratio of initial estimated ultimates (at end of the first development year) to the latest estimated ultimates (at end of 2001) for each of the previous 20 accident years. When plotted against time, this ratio appeared, in both cases, to show a cyclical pattern of under- and over-reserving. This cycle appeared to be in phase with the underwriting cycle over the nearly two complete cycles of the years 1980 to 2001. Initial reserve estimates were consistently too low in both of the "soft market" periods when premium rates were relatively low (the mid 1980s and the late 1990s), and were consistently too high in the intervening "hard markets" (when premium rates were relatively high).

The most obvious explanation is that when setting reserves soon after writing the business, insurers tend to under-estimate the magnitude of the underwriting cycle. History shows that at the lowest point of the underwriting cycle, insurers often write business at loss-making rates. But presumably they don't do this deliberately: it is only human to hope and believe that the business they recently wrote will ultimately prove profitable, and to set the initial reserves accordingly. At the other extreme of the underwriting cycle, when premium rates are buoyant, even relatively cautious (high) initial reserves may show a fairly good profit. Management might privately believe that they could reasonably set the initial reserves lower and show an even higher profit. But with a choice between declaring a very high profit now, with the possibility that this will deteriorate, and declaring a more moderate (but still healthy) profit now, with the expectation that this will allow further good news to be released as the claims run-off, it is easy to see the attraction of the latter.

If this were the whole explanation for the reserving cycle, then the actuarial profession could rest easy. If we as actuaries provide objective, unbiased estimates for the reserves, and senior management chooses to depart from these estimates for reasons such as those described above, then that is their responsibility not ours.

However, we need to be sure that actuarial reserve estimates are as good as they can be. Could it be that actuarial reserving methods are partly to blame for the reserving cycle?

1.1.2 Working party report at GIRO 2003

Bob Conger's presentation at GIRO 2002 prompted the formation of a working party tasked with investigating the existence and possible causes of a reserving cycle in the UK. This working party was chaired by Nick Line, and presented its report [3] at the 2003 GIRO convention. The working party concluded that:

(a) A reserving cycle did also exist in the UK.

(b) Standard actuarial reserving methods are probably a contributory cause of the reserving cycle.

(c) There was some (inconclusive) evidence that development patterns vary with the underwriting cycle, tending to be longer-tailed when premium rates are low.

(d) There was clear evidence that Lloyd's premium rate indices had tended to understate the true magnitude of the underwriting cycle.

Conclusion (a) is based on UK industry reserves (from regulatory returns) over the period 1985 to 2001 inclusive. This is a shorter period than that considered by Bob Conger, and covers little more than one complete underwriting cycle (including the soft market of the mid to late 1980s and the next soft market of the late 1990s). The working party looked at the non-life insurance market as a whole, and some individual classes, and concluded that the reserving cycle exists for several major classes (motor, property and liability) and that the cycles for these classes are in phase with one another.

The reserves analyzed by the working party in support of conclusion (a) were obtained from regulatory returns, and as such show booked reserves as opposed to actuarial estimates. Conclusion (b) was based on an investigation of the extent to which standard actuarial reserving methods (chain ladder (CL) and Bornheutter-Ferguson (BF)) produce cyclical under- and over-reserving if applied mechanistically. This was investigated by applying these methods to the run-off data from regulatory returns, and comparing early estimates to actual ultimates. For some (but not all) classes of insurance, the results showed a clear cyclical pattern closely following that observed in the regulatory reserves. This was clearest for long-tail liability classes.

The working party then tried to explain why these standard actuarial methods tend to give a cyclical pattern of under- and over-reserving. They postulated two main causes: the points labeled (c) and (d) above. The apparent variation in development patterns with the underwriting cycle (point (c)) violates the basic assumption of the CL and BF methods: that the development pattern is the same for all origin periods. The working party found some evidence of more rapid paid development for origin years in the "hard" part of the underwriting cycle. The paid chain ladder method would clearly tend to overstate reserves at the top of the cycle and understate at the bottom of the cycle if this is the case.

The tendency for premium-rate indices to understate the amplitude of the underwriting cycle (point (d) above) exacerbates the reserving error produced by the BF method: if the softness of a soft market is understated, then the prior expected loss ratio will also be understated, leading to an initial under-estimation of the ultimate.

[Note that if both (c) and (d) apply, their effects on paid BF reserves will be in the same direction, rather than offsetting each other. The paid BF reserve is (1-F)*(prior ultimate), where F is the expected proportion of ultimate development obtained by the chain ladder method. The CL method gives a value for F that is an average for all origin years in the run-off array. If development is quicker than average when premium rates are high (point (c)), then this will be lower than the expected proportion developed for accident years where premium rates are high, so the factor (1-F) will be too high for these years. Assuming the other factor of the BF reserve (the prior ultimate) is calculated in the usual way (as premium multiplied by average ULR divided by premium index) this factor will be too high if the premium index is too low when premium rates are high (point (d)).

Conversely, both factors of the paid BF reserve will be too low for origin years where premium rates are low.]

The 2003 GIRO working party also suggested some possible reasons why the run-off pattern might depend on the level of premiums (that is, possible explanations for point (c)). They came up with the following possible causes of longer paid development patterns in soft markets. (Note that the working party did not look for direct evidence that any of these actually occur: these points were merely suggested as possible causes.)

- When premium rates are low, insurers might be more reluctant to pay claims, leading to more protracted negotiations and longer payment delays.
- In soft markets, insurers might compete by including additional cover and relaxing terms and conditions (as well as by reducing premiums). This might result in more disputes over coverage, tending to lengthen development patterns. (Presumably tightening of terms and conditions in hard markets might equally lead to disputes, but the difference is that in this case, disputes are less likely to delay payments.)
- If upper policy limits are increased in soft markets, this would also tend to lengthen development patterns. (On the other hand, if insurers reduce deductibles in soft markets, this would tend to shorten development patterns because deductibles would be exhausted sooner.)
- If more multi-year policies are written in soft-markets, these would tend to lengthen development patterns of under-writing year cohorts (but this should not affect accident year run-off patterns).

The above four points relate to paid development. In addition, the working party noted that incurred development patterns would be longer-tailed in soft markets if insurers adopt more optimistic case-reserving practices when premium rates are low.

1.1.3 General Insurance Reserving Issues Taskforce (GRIT)

At the beginning of 2004, the UK actuarial profession created the General Insurance Reserving Issues Taskforce (GRIT). Of five specific issues given in the terms of reference for GRIT, one was "to consider the actions which the profession should take in relation to the observations made in the Reserving Cycle Working Party paper presented at GIRO 2003." GRIT produced its final report (after a consultation process within the UK profession) in March 2006 [2]. Section 7 of the GRIT report, entitled "Improving our Methods," is mostly concerned with reserving cycles. This topic is also mentioned in Sections 1.1.6, 1.1.7, 1.8.2-1.8.5, 2.6.4, 9.2 and 9.9 of the GRIT report.

GRIT carried out basically the same analysis as the 2003 working party, but using Lloyd's data where the previous working party had used insurance company data (from regulatory returns). Like the working party, GRIT applied the CL and BF methods mechanically to historical run-off triangles for different classes of business, and compared early forecasts produced by these methods to actual outcomes. Like the working party, GRIT concluded that these methods do produce a cyclical pattern of under and over reserving, and that this pattern is in phase with the underwriting cycle.

As a possible way forward, GRIT suggested (Section 7.5 of [2]) fitting cumulative Weibull distribution curves to cumulative paid development data, and allowing the scale parameter to vary cyclically. The equation for the cumulative development pattern proposed in [2] is:

$$\text{Claims}(t) = A * [\ 1 - \exp\{-(b/t)^c\}] \tag{1}$$

Here, t is development time, A is ultimate, b is a scale parameter, and c is a shape parameter. [Note that c has to be negative in order for this to be a valid cumulative development pattern: if c is positive, then Claims(t) tends to zero as t tends to infinity. The Weibull curve is usually specified using t/b instead of b/t so that the shape parameter c takes positive values: we then have Claims(t) tending to the ultimate A as t tends to infinity.]

1.1.4 Other prior research on the reserving cycle

In the UK, GRIT was replaced (following publication of its final report in 2006) by the General Insurance Reserving Oversight Committee (GI ROC). GI ROC initiated four working parties that would report to future GIRO conventions. One of these is the working party on "Implications of the underwriting and reserving cycles for reserving." By the time of GIRO 2007 this working party had not made significant progress.

Perhaps surprisingly, considering that it was Bob Conger (then president of the CAS) who first highlighted this issue, there seems to be no published research in this area by US actuaries. (At least, a search of the CAS Web Site yields nothing new.)

1.2 Objectives of the Present Paper

The present paper develops the idea (suggested in the 2006 GRIT report [2]) of fitting curves to cumulative development data in a way that allows for the possibility of cyclical variation of development patterns.

The GRIT paper suggested using a cumulative Weibull distribution function (see Equation 1) for this purpose. In the present paper, Weibull, Burr, and Inverse Burr distribution functions are used.

Any distribution function is by definition an increasing function. In practice, cumulative incurred run-off patterns often do not increase at all stages of development, so cumulative probability distribution functions would not provide a good fit. The present paper develops a family of curves that does have the flexibility to accommodate typical cumulative incurred development patterns. This family of curves is derived by modeling both reporting and payment delays using cumulative probability distribution functions. This produces two linked families of cumulative curves: one for paid, the other for incurred. By fitting these simultaneously to paid and incurred run-off data, a single ultimate is estimated for each origin year from all available data. This avoids the common problem of having one ultimate estimated from paid data and another ultimate estimated from incurred data, then having to combine the two somehow.

Parameters of the fitted curves are linked to a premium rating index so that both paid and incurred run-off curves are allowed to vary with the underwriting cycle. This is done in a way that allows for the possibility that the premium rating index might understate the true amplitude of the underwriting cycle (as found to be the case in [3]).

The paper is not concerned with directly looking for evidence of each the possible causes of cyclically varying run-off patterns discussed in Section 1.1.2. Instead, the mathematical model is developed in such a way that it will accommodate these possible causes if they exist. The model also accommodates other possible factors, such as variation in reporting delay with the underwriting cycle. It will not always be possible to distinguish the true cause using the results of fitting the proposed model.

2. CYCLICAL CURVE-FITTING METHOD

2.1 Principles of curve-fitting to claims development data

In Section 2, a model is introduced that can be used to test for the presence of cyclical development patterns, to distinguish some of the main cyclical effects, and to estimate ultimates in the presence of these effects. Initially it is assumed that the only data available are the usual aggregate cumulative paid and incurred run-off arrays, and a premium rate index. Later (Section 2.4) the use of premium or other exposure information is also considered.

The method is basically a curve-fitting method as suggested in the GRIT report [2]. When curve-fitting is used for reserving, it is common to assume that the run-off pattern is the same for all origin years. If this is not the case, origin years can sometimes be grouped so that it is approximately the case in each group. The GRIT paper suggested classifying origin years into two categories according to their position in the underwriting cycle (hard or soft). Instead of doing this, the method introduced in the present paper uses a premium rate index to allow continuous graduation between hard and soft market run-off curves.

Reserving methods have previously been developed that allow run-off patterns to gradually change across origin years: for example, the method described in Wright [4]. That method allows for trend changes in development patterns but not for cyclical changes. It is also quite complex because it is a full stochastic method which gives predictive standard errors as well as best estimates. A limitation of that method is that it requires mainly positive increments in the run-off data, so it often cannot be applied directly to incurred data without first adjusting the data in some way.

What we need now is a method that can be applied to both paid and incurred data, preferably making use of premium development data too, and which allows for cyclical changes in run-off patterns. The top priority is to develop such a method that gives good point estimates in the presence of underwriting and reserving cycles. A lower priority is rigorous assessment of standard errors: this is not considered in the present paper.

If the run-off pattern is not assumed to be the same for all origin years, then the model necessarily has more parameters than where the run-off pattern is assumed to be constant. It is a well-established statistical principle that as the number of estimated parameters increases, their reliability (when estimated from a given volume of data) generally decreases. Therefore, it is advisable to use as much relevant data as possible when estimating parameters. For this reason, the proposed method fits run-off curves to both paid and incurred data simultaneously. This also avoids the problem (met with most other reserving methods) of having one set of reserve estimates obtained from paid data and a different set of estimates obtained from incurred data, then having to combine into a single set of final estimates. In order to allow fitting to incurred data as well as paid, the family of run-off curves must allow for negative increments as these often exist in incurred data.

2.2 Model for a single origin year

2.2.1 Cumulative paid and incurred development curves

In this sub-section we consider the run-off of a single origin year. The model is generalized for multiple origin years in later sub-sections.

$F_P(t)$ is used to denote a cumulative paid run-off pattern (where t is continuous development time). This is a function that starts at 0 when $t = 0$, and increases to 1 as t tends to infinity. For paid data, although there may be occasional decreases due to salvage and subrogation, the underlying pattern is assumed to be strictly increasing. $F_P(t)$ is therefore a cumulative distribution function: its derivate $f_P(t) = dF_P(t)/dt$, can be regarded as the probability density function for the delay to payment of each dollar that is ultimately paid.

For modeling incurred run-off patterns, we need to consider reporting delays. We use $F_R(t)$ to denote the cumulative distribution function of reporting delays (in respect of each dollar that is ultimately paid). Since every claim must be reported before it is paid, we should have $F_R(t) \geq F_P(t)$ at all development times t. Exhibit 1 shows typical curves $F_R(t)$ and $F_P(t)$ for a single origin year. (The curves in Exhibit 1 are Weibull distributions with mean values of 1 year and 3.6 years respectively.)

Note that $F_R(t)$ is not the cumulative incurred development pattern: it is the distribution function of reporting delays in respect of each dollar that is ultimately paid. For example, consider an accident year with an ultimate paid amount of $100,000. Suppose the first claim is reported mid-way through accident year zero (at time $t = 0.5$), and that this claim ultimately settles for $1,000. Since this is 1% of the total ultimate for the accident year, $F_R(t)$ increases from 0 to 0.01 at $t = 0.5$. The incurred development pattern will usually differ from this. For example, suppose that when this first claim is reported the initial case reserve is set at $2,000. Since this is 2% of the total ultimate for the accident year, the incurred development pattern increases from 0 to 0.02 at $t = 0.5$. (Of course, none of these development patterns is known with certainty until the accident year concerned is fully developed.)

To model incurred development, we assume that when a claim is reported a case reserve is set up, and the amount of the case reserve (on average, in the period between reporting and eventual payment) is b-dollars for each dollar that is ultimately paid. If case reserves are set conservatively (perhaps more likely during hard markets) we will have $b > 1$. In soft markets, case reserves are more likely to be set optimistically so b may take lower values, and we might have $b < 1$.

Under the above assumptions, for each dollar of ultimate, the expected cumulative amount paid by development time t is $F_P(t)$ and the expected amount outstanding at time t is $b.\{F_R(t) - F_P(t)\}$. So if we use $F_I(t)$ to denote the expected cumulative incurred run-off pattern then (from the definition of incurred as paid plus outstanding) we have:

$$\begin{aligned} F_I(t) &= F_P(t) + b.\{ F_R(t) - F_P(t)\} \\ &= b.\, F_R(t) + (1- b).\, F_P(t). \end{aligned} \tag{2}$$

This last equation can be interpreted by noting that incurred increases by the amount b when the claim is reported (and the case reserve set up), then increases by the amount (1- b) (which is usually negative) when the claim is paid.

Because of the possibility that $b > 1$, the function $F_I(t)$ is not in general a probability distribution function because it is not strictly increasing. Both $F_R(t)$ and $F_P(t)$ are strictly increasing (from 0 to 1) but if $b > 1$, $F_I(t)$ will show the usual incurred run-off shape: increasing rapidly then decreasing towards ultimate. This is illustrated in Exhibit 2, which shows typical run-off patterns for the case $b = 1.5$ (that is, case reserves are on average 50% higher than what is ultimately paid in respect of the reported claims).

Suppose we have cumulative paid and cumulative incurred run-off data. If we assume some parametric family of curves for $F_P(t)$ and $F_R(t)$, Equation 2 then implies a parametric family for $F_I(t)$. The parameters can be estimated by fitting the curve $F_P(t)$ to the cumulative paid data, and the curve $F_I(t)$ to the incurred data. Note that b is one of the parameters that will be estimated from the data. Some suitable parametric distributions for $F_P(t)$ and $F_R(t)$ are considered in the next sub-section.

The bias factor b need not necessarily be assumed to take a constant value across development time (within each single origin year). It seems likely that the accuracy of case reserves might sometimes improve with time. This possibility can be allowed for by using a model of the form:

$$b_t = \exp\{\beta_0 + \beta_2.\max(0, t_0\text{-}t)\}/ \tag{2a}$$

Instead of a single constant parameter b, this form of model has three parameters β_0, β_2 and t_0. (A further parameter, β_1, is introduced in Section 2.3.2.) The exponentiation ensures that the bias factor b_t is always positive. The expression max(0, t_0-t) allows b_t to change between development times t=0 and t=t_0. At later development times, max(0, t_0-t) is zero so the bias factor b_t settles at $\exp\{\beta_0\}$.

2.2.2 Suitable parametric distribution families

When selecting a family of curves to fit to paid claims development data, in principle any analytic family of probability distribution functions could be tried: for example Log-Normal, Pareto, Gamma, Weibull, etc. However, in this paper we restrict attention to distribution families that have the following properties:

- The cumulative distribution function is mathematically simple. This is desirable so that curve-fitting can be carried out quickly and easily in a spreadsheet.
- Ability to accommodate a wide range of values for the ratio of mode to mean. In particular, the distribution family should include distributions with mode equal to zero as well as distributions with mode greater than zero. [Recall that the mode of a distribution is the point where the density function takes its maximum value, or equivalently, where the slope of the cumulative distribution function is greatest.] This is desirable so that the same family of distributions can reasonably be used for reporting delays and for payment delays (which is merely convenient, not strictly necessary). In some classes of insurance, reporting delays tend to be very short for the majority of claims so that the mode is close to zero. For payment delays, the mode is invariably greater than zero.

The Log-Normal distribution (for example) is not used in this paper because it does not satisfy either of these criteria. The Log-Normal cumulative distribution function can be calculated from the Normal distribution function (which is available in popular spreadsheet software) but it is relatively complex and slow to calculate compared to some simpler distribution functions. The mode and mean of a Log-Normal distribution (using the usual μ and σ parameterization) are respectively $\exp(\mu)$ and $\exp(\mu + \sigma^2/2)$, so the ratio of mode to mean is $\exp(-\sigma^2/2)$. Although this can take any value between zero and one, a mode of zero is not possible.

Based on the above criteria, three distribution families have been selected for use in the present paper to model development patterns. Other distribution families could be used within the framework developed here, and some might prove to be more suitable than the selected three. These three distribution families have been chosen for convenience, on the basis that we have to start somewhere, and because they are probably as good as any for illustrating the principles of the proposed method.

The three distribution families used in this paper are the Weibull, the Burr, and the Inverse Burr. Table 1 gives the cumulative distribution function $F(t)$ and the mean and the mode of these distributions. (The penultimate column gives conditions for the mode to be greater than zero, and the final column gives the formula for the mode when it is greater than zero.) $\Gamma(.)$ denotes the Gamma function, which can be evaluated in Excel® as EXP(GAMMALN(x)).

Table 1: Formulas for analytic delay distributions

	$F(t)$	Mean	Mode>0 if	Mode (if > 0)
Weibull	$1 - \exp\{-(t/s)^c\}$	$s.\Gamma(1+1/c)$	$c > 1$	$s.(1-1/c)^{1/c}$
Burr	$1 - 1/\{1+(t/s)^c\}^a$	$s.\Gamma(1+1/c).\Gamma(a-1/c) \;/\; \Gamma(a)$	c > 1	$s.\{(c-1).(ac+1)\}^{1/c}$
Inv Burr	$1/\{1+(s/t)^c\}^a$	$s.\Gamma(a+1/c).\Gamma(1-1/c) \;/\; \Gamma(a)$	$ac > 1$	$s.\{(ac-1).(c+1)\}^{1/c}$

In all three cases, $F(t)$ increases monotonically from 0 when t=0, towards 1 as t tends to infinity. The parameter s is a scale parameter; the parameters a and c are shape parameters. The Weibull family has just one shape parameter; the Burr and Inverse Burr families each have two shape parameters. The additional shape parameter means that the Burr and Inverse Burr families are much larger and more flexible then the Weibull family. The Burr and Inverse Burr families have some well-known sub-families. The Pareto is the sub-family of the Burr family obtained by setting c to 1. The Inverse Pareto is the sub-family of the Inverse Burr obtained by setting c to 1. The Log-Logistic is the sub-family of both Burr and Inverse Burr families obtained by setting a to 1. Each of these sub-families has one shape parameter, so in that sense, is as large as the Weibull family.

Although we use all three of these distribution families in Section 3 of this paper, in the remainder of Section 2 we use the Weibull distribution for both payment delays and reporting delays. The Weibull is used because it is a relatively simple distribution, and it serves to illustrate the principles of the proposed modeling method. No implication is intended that the Weibull is superior to any other distribution family for this purpose. As already noted, any analytic distribution family could be used within the framework developed in Section 2.2.1, and using the same principles as are illustrated below using the Weibull distribution. There is also no reason in principle why two different distribution families should not be used; one to model reporting delays $F_R(t)$, and another to model payment delays $F_P(t)$, (provided $F_R(t) \geq F_P(t)$ for all values of t).

2.2.3 Model for single origin year based on Weibull distributions

Here the Weibull distribution is used to model both reporting and payment delays. Subscripts R and P are used to distinguish parameters of the reporting and payment delay distributions. The symbol ^ is used to denote exponentiation, that is: (t/s)^$c = (t/s)^c$.

For the reporting and payment delay distributions we have:

$$F_R(t) = 1 - \exp\{-(t/s_R)\hat{}c_R\} \tag{3}$$
$$F_P(t) = 1 - \exp\{-(t/s_P)\hat{}c_P\} \tag{4}$$

So from Equation 2, the incurred development pattern is:

$$F_I(t) = 1 - b.\exp\{-(t/s_R)\text{^}c_R\} - (1\text{-}b).\exp\{-(t/s_P)\text{^}c_P\} \tag{5}$$

The above equations for $F_P(t)$ and $F_I(t)$ specify a model for the paid and incurred run-off patterns for each dollar of ultimate. If U denotes the ultimate paid (which is what we aim to estimate), then the expected amounts paid and incurred by development time t are respectively $U.F_P(t)$ and $U.F_I(t)$.

From Equations 3 and 4, the requirement $F_R(t) \geq F_P(t)$ (for all values of t) is equivalent to $(t/s_R)\text{^}c_R \geq (t/s_P)\text{^}c_P$, which in turn is equivalent to $t\text{^}(c_R\text{-}\ c_P) \geq (s_R\text{^}c_R)/(s_P\text{^}c_P)$. If c_R is not equal to c_P, then the left side of this inequality takes all values between zero and infinity as t varies between zero and infinity. So the only way this inequality can be true for all positive values of t is by having c_R equal to c_P (so the left side is equal to 1 for all t) and s_R less than s_P (so the right side is less than 1). However, in practice, it is of little consequence if $F_R(t)$ is less than $F_P(t)$ for high values of t (that is, where both $F_R(t)$ and $F_P(t)$ are very close to 1). So we will not insist on the constraint $c_R = c_P$. Instead, it is proposed to check that the fitted curves are reasonable by viewing them graphically. This is illustrated by Exhibit 1, in which both curves are Weibull distributions, with parameters $s_R = 1$, $c_R = 1$, $s_P = 4$, $c_P = 3$. These parameters give $F_R(t) \geq F_P(t)$ for $t \leq 8$, but $F_R(t) < F_P(t)$ for $t > 8$. Since both curves reach 99.97% development at $t = 8$, it is of no practical consequence that $F_R(t) < F_P(t)$ for $t > 8$.

2.3 Model for multiple origin years

2.3.1 Variation of parameters across origin years

In Section 2.2 we developed a model for paid and incurred development patterns of a single origin year. If the Weibull distribution is used for both reporting and payment delays (as in 2.2.3), then the model has six parameters for each origin year: U, b, s_R, c_R, s_P, c_P. This is clearly too many parameters to attempt to estimate separately for each origin year. For the latest origin year we usually have only two data values: one paid, one incurred (although there may be more if sub-annual development periods are used).

The total number of parameters needs to be reduced. To achieve this, we could try assuming initially that some of the parameters take a single constant value across all origin years. For example, we might assume that s_R, c_R and c_P take the same values for all origin years, so that only the parameters U, b and s_P vary across origin years. Setting s_R and c_R to values that are constant across all origin years is appropriate if reporting delays have the same distribution $F_R(t)$ across all origin years. For some datasets this might turn out to be a reasonable assumption. (We discuss later how this assumption can be tested. Applications of the model presented in this paper to actual datasets have shown evidence that reporting delays do sometimes vary with the underwriting cycle: possible causes are discussed in Section 4.1.3.)

By allowing b to vary across origin years, we allow for the possibility that case reserves are set up more or less conservatively at different points in the underwriting cycle. This possibility is suggested in the existing literature discussed in Section 1. By allowing the scale parameter s_P of the payment delay distribution to vary, we allow for variation in the speed of claim settlement. Previous research has found evidence that such variation does occur with the underwriting cycle (see Section 1). Clearly, the ultimate U must also be allowed to take a different value for each origin year as this is what we aim to estimate.

In the remainder of this paper, parameters that are allowed to vary across origin years have a subscript j to label the origin year. So in the Weibull model, if parameters U, b and s_P are allowed to vary, these are denoted U_j, b_j, s_{Pj}.

2.3.2 Allowing for cyclical development patterns

To allow for underwriting cycle effects, parameters of the run-off curves that are not held constant across all origin years can be linked to a known premium rate index. Interpretation of model parameters is simplified if the premium rate index (denoted Q_j for origin year j) is scaled so the mean value across all origin years is 1. We can then use equations of the form:

$$b_j = \exp\{\beta_0 + \beta_1.(Q_j\text{-}1)\} \qquad (6)$$

$$s_{Pj} = \exp\{\sigma_0 + \sigma_1.(Q_j\text{-}1)\} \qquad (7)$$

Here, β_0, β_1, σ_0 and σ_1 are parameters (to be estimated from the run-off data) that are assumed to take the same values for all origin years. The subtraction of 1 from the premium rate index further simplifies interpretation of the parameters: for example, $\exp\{\beta_0\}$ represents the value of b for an average year in which $Q_j = 1$. The exponentiation ensures that parameters b_j and s_{Pj} are always positive (which is necessary to produce valid development curves). Note that this form of model for b_j and s_{Pj} allows for the possibility that the known index Q_j may understate the true amplitude of the underwriting cycle. Previous research (see Section 1) suggests that this is often the case with premium rate indices. If this is the case, the parameters β_1 and σ_1 estimated from the run-off data will simply take higher values than they would if Q_j correctly reflected the amplitude of the reserving cycle. The number of parameters to be estimated from the paid and incurred run-off data is now reasonable. We have:

- Seven parameters each assumed to take a single constant value across all origin years (s_R, c_R, c_P, β_0, β_1, σ_0 and σ_1), plus
- One parameter taking a different value for each origin year (the ultimate U_j).

If there are more than eight development periods, the basic chain ladder model has more parameters than this.

If the possibility that bias factors change with development time is allowed for as described in Section 2.2.1 (Equation (2a)) then Equation (6) becomes:

$$b_t = \exp\{\beta_0 + \beta_1.(Q_j\text{-}1) + \beta_2.\max(0, t_0\text{-}t)\} \tag{6a}$$

2.3.3 Estimation of parameters by least squares

The parameters of the paid and incurred development curves can be determined by the method of least-squares. The following notation is used in this section:

- $F_{Pj}(t)$ denotes the cumulative paid development curve for origin year j. Using the Weibull model, this is given by Equation 4, with scale parameter s_P replaced by s_{Pj} from Equation 7.
- $F_{Ij}(t)$ denotes the cumulative incurred development curve for origin year j. Using the Weibull model, this is given by Equation 5, with scale parameter s_P replaced by s_{Pj} from Equation 7, and the case-reserve redundancy-factor b replaced by b_j from Equation 6 (or Equation 6a).
- P_{jd} denotes the actual cumulative paid for origin year j and development period d.
- I_{jd} denotes the actual cumulative incurred for origin year j and development period d.

The residual sum of squares is defined as the sum of squared differences between actual and expected values. This can be calculated separately for paid and incurred:

$$\begin{aligned}\text{RSS}_P &= \Sigma\ \{P_{jd} - U_j.F_{Pj}(t)\}^2 \\ \text{RSS}_I &= \Sigma\ \{I_{jd} - U_j.F_{Ij}(t)\}^2\end{aligned} \tag{8}$$

Summation is over all origin years j and development periods d in the run-off arrays. In the case of annual development data, d denotes the development year. We use the convention that the origin year itself is development year 0, so d takes the values 0, 1, 2, etc. In the fitted curves ($F_{Pj}(t)$ and $F_{Ij}(t)$), t denotes continuous development time. Ideally this would be the exact elapsed time from the date of loss occurrence. However, since claims in a particular origin year cohort do not usually all have exactly the same date of loss occurrence, t is set to an approximate average delay from the date of loss occurrence until the end of the corresponding development period d. Table 2 gives appropriate values of t for each development year d, for both accident year and underwriting year cohorts. These approximations are based on assumptions that accidents occur uniformly in time and policies incept uniformly. Further details are given in Appendix B.

Similar approximations can be used for sub-annual development periods. Approximations such as these tend to be relatively crude for early development periods: this is discussed further in Section 2.5.3.

Table 2: Approximate mean delay in each development year

Development year (d)	0	1	2	3	4	5+
Accident year mean delay (t)	0.5	1.5	2.5	3.5	4.5	d+0.5
Underwriting year mean delay (t)	0.333	1	2	3	4	d

Given values for the parameters of the development curves (in the case of the Weibull model: s_R, c_R, c_P, β_0, β_1, σ_0 and σ_1) and a value for the ultimate U_j of each origin year, the "expected" values $U_j.F_{Pj}(t)$ and $U_j.F_{Ij}(t)$ can be calculated corresponding to each cell (j,d) of the run-off array. From these, the residual sums of squares RSS_P and RSS_I can be calculated (Equation 8). The least squares estimation method is to search for the values of the parameters (s_R, c_R, c_P, β_0, β_1, σ_0, σ_1 and U_j for each origin year) that minimize the residual sums of squares. Note that ultimates U_j can be treated as parameters of the model and estimated by least squares along with the other parameters. However for early origin years, the ultimate may already be known with some precision. If, for a particular origin year all reported claims have been settled and further claims are considered unlikely, then there is no need to estimate the ultimate U_j by least squares, and a better model will usually be obtained by using the known value of this quantity. Usually this applies only for the earlier origin years: for origin years that are not fully developed the ultimate is estimated by least squares.

2.3.4 Combining paid and incurred by weighted least squares

It is clearly possible (provided the number of data-points exceeds the number of parameters) to carry out least squares estimation separately for paid and incurred. However, the paid and incurred models have parameters in common. (In the case of the Weibull model of Section 2.3.2, the following parameters feature in both the paid and incurred models: c_p, σ_0, σ_1 and the ultimate U_j of each origin year.) If least squares estimation is carried out separately for paid and incurred, the paid data will yield one set of estimates for these parameters, and the incurred data will yield another set of values for the same parameters. This can be avoided by carrying out the least squares procedure just once based on the total residual sum of squares $RSS_P + RSS_I$.

This raises the question of relative weighting between paid and incurred: is it correct to give RSS_P and RSS_I equal weight by just adding them? An alternative would be to find the parameter values that minimize $RSS_P+w.RSS_I$ where w is a predefined weighting factor. To increase the influence of the incurred data relative to the paid data, we would choose a value for w that is greater than 1, and to give more influence to the paid data we would choose a value less than 1.

One way to justify a relative weighting on theoretical grounds would be to develop a full stochastic model that treats each value P_{jd} and I_{jd} as a random variable and gives an expression for the variance of each one. The basic theoretical justification for the least squares method is two-fold:

- The Gauss-Markov theorem states that, in linear models, weighted least squares estimates have the smallest variance of all linear unbiased estimates.
- Quasi-likelihood theory shows that weighted least squares estimates are asymptotically unbiased and efficient (that is, have minimum possible variance) even in non-linear models.

In both Gauss-Markov and quasi-likelihood theory, the weights that give optimal least squares estimates are inversely proportional to the variances of the corresponding random variables. In addition, if some of the random variables are correlated, then the residual sum of squares that is minimized should include cross terms with weights depending on the covariance between the corresponding variables.

In the present application, it is clear that the observations (P_{jd} and I_{jd}) are not all mutually independent. Since incurred is paid plus outstanding, any reasonable stochastic model would indicate a positive covariance between the values P_{jd}, I_{jd} with the same values of j and d. Furthermore, since these are cumulative values, it is likely that there is serial correlation between successive values of P_{jd} as the development period d increases in each fixed origin year j. For this reason, a full stochastic model would indicate that the optimum "residual sum of squares" to be minimized should include cross terms such as $\{P_{jd} - U_j.F_{Pj}(t)\}.\{I_{jd} - U_j.F_{Ij}(t)\}$ as well as pure squared terms such as $\{P_{jd} - U_j.F_{Pj}(t)\}^2$. A full stochastic model would also indicate the optimum relative weighting of every term in the residual sum of squares.

Development of a full stochastic model is not attempted in this paper because it would be mathematically complex and probably contentious (as it would require many assumptions about the nature of the stochastic variation in paid and incurred run-off data). Instead, the aim is to produce a method that is simple enough to be widely useful if applied intelligently. To this end, it is proposed to ignore the clear correlation that will exist between the observed values of P_{jd} and I_{jd} by including no cross-terms in the residual sum of squares. We also take no account of differing variances among the P_{jd} (and among the I_{jd}) by giving every term equal weight in RSS_P (and in RSS_I). The remaining question is: should we take this cavalier approach one step further by giving equal weight to both RSS_P and RSS_I ?

2.3.5 Empirical determination of relative weighting of paid and incurred

At this point, it is proposed to allow for the possibility that one of the two datasets (either paid or incurred) may appear to be more reliable than the other. The theory (Gauss-Markov and quasi-likelihood) suggests that the two terms (RSS_P and RSS_I) should be weighted in inverse proportion to the mean variance of paid and incurred data-points. That is, instead of minimizing RSS_P+RSS_I we should minimize $(\text{RSS}_P/\sigma^2_P)+(\text{RSS}_I/\sigma^2_I)$, where σ^2_P and σ^2_I are typical variances of individual paid and incurred observations. This is equivalent to minimizing the following total weighted sum of squares:

$$\text{Weighted sum of squares} = \text{RSS}_P + w_I.\text{RSS}_I \text{ where } w_I = \sigma^2_P/\sigma^2_I. \tag{9}$$

Instead of using a stochastic model to determine the relative magnitudes of σ^2_P and σ^2_I, a purely empirical approach can be used in which their relative magnitudes are estimated from the residuals. Standard theory suggests the variance of a paid observation be estimated as:

$$\sigma^2_P = \text{RSS}_P/(n_P - p_P) \tag{10}$$

where n_P is the number of paid observations, and p_P is the number of parameters estimated from these observations. If variances are estimated in this way we will have $(\text{RSS}_P/\sigma^2_P) = (n_P - p_P)$ and $(\text{RSS}_I/\sigma^2_I) = (n_I - p_I)$, so the total weighted sum of squares $(\text{RSS}_P/\sigma^2_P)+(\text{RSS}_I/\sigma^2_I)$ will be $(n_P + n_I) - (p_P + p_I)$. An iterative fitting procedure is necessary to achieve this. It is also necessary to divide the total parameter count into the two components p_P and p_I. Each parameter that features in the fitted curves of both paid and incurred (for example, the ultimates U_j) makes a fractional contribution to both p_P and p_I. What is believed to be a reasonable pragmatic approach is proposed for this purpose. (This is discussed further in Section 2.5.1, but no rigorous theoretical justification is claimed.)

For example, suppose we have annual paid and incurred run-off arrays for 10 origin years, so that $n_P = n_I = 55$. Suppose we are using the Weibull model described in Section 2.3.2. The paid development curves depend on 13 parameters: c_P, σ_0, σ_1 and U_1, $U_2 \ldots U_{10}$. The incurred development curves depend on 17 parameters: the same 13 as in the paid model, plus s_R, c_R, β_0 and β_1. If we count a parameter that features in both paid and incurred models as half a parameter in each model, we have: $p_P = 6.5$ and $p_I = 10.5$, which gives the correct total number of distinct parameters: $p_P + p_I = 17$. (A more refined method of counting parameters is discussed in Section 2.5.1.)

On the first iteration, we estimate the parameters by minimizing RSS_P+RSS_I (that is, using equal weights initially). Suppose this produces parameter values that give $RSS_P = 100$ and $RSS_I = 300$, so the total minimized residual sum of squares is 400. Initial estimates of the mean variances for paid and incurred are then: $\sigma^2_P = 100/(55–6.5) = 2.06$ and $\sigma^2_I = 300/(55–10.5) = 6.74$ (If we had reason to believe that σ^2_P and σ^2_I were equal, we would estimate the value as 400/(110–17) = 4.30.) σ^2_I being so much higher than σ^2_P indicates that the model does not fit the incurred data as closely as it fits the paid data. The incurred data are therefore less reliable than the paid for the purpose of projecting run-off patterns, and so should be given less weight than paid in fitting the model. So for the next iteration, instead of minimizing RSS_P+RSS_I, we minimize (RSS_P/2.06)+(RSS_I/6.74). Multiplying by 2.06, we see that this is equivalent to minimizing $RSS_P + 0.31 * RSS_I$. Minimizing this might result in $RSS_P = 95$ (that is, a closer fit to the paid data than on the first iteration) and $RSS_I = 350$ (a poorer fit to the incurred data than on the first iteration). These figures give the following revised estimates of variances: $\sigma^2_P = 95/(55–6.5) = 1.96$ and $\sigma^2_I = 305/(55–10.5) = 7.87$. So on the third iteration, we minimize $RSS_P + 0.25 * RSS_I$. Continuing in this way, convergence usually occurs after a few iterations.

If the model is set up in Excel®, the Excel solver can be used to search for the parameter values that minimize the required weighted sum of squares in each iteration. Note however, that since the fitted curves are non-linear functions of the parameters, solver does not guarantee to find the global minimum. It is advisable to try several sets of starting values if there is any doubt about the solution found by the Excel solver.

2.4 Use of premium and exposure data

2.4.1 Estimated ultimate for latest origin year

In the model described so far, no use has been made of premium or other exposure data. For the latest origin year ($j = J$ say), the ultimate U_J is estimated purely by fitting curves to development patterns, and assuming that changes in the parameters of these curves from one origin year to the next are linked to changes in the underwriting cycle (through equations such as 6 and 7). The latest origin year has one free parameter of its own (the ultimate U_J), and (assuming annual paid and incurred development data) two data-values: the actual paid and incurred amounts at the end of the zeroth development year P_{J0} and I_{J0}.

This latest origin year contributes just two terms to the weighted sum of squares:

$$\{P_{J0} – U_J.F_{PJ}(t)\}^2 + \mathrm{w}.\{I_{J0} – U_J.F_{IJ}(t)\}^2 \tag{11}$$

Here, t = ½ for accident year cohorts, or 1/3 for underwriting year cohorts: see Table 2.

Since U_J does not appear in any other terms of the total sum of squares, it can be adjusted to minimize the sum of the above two terms. Elementary calculus shows that this gives:

$$U_J = \{P_{J0}.F_{PJ}(t) + w.I_{J0}.F_{IJ}(t)\} / \{F_{PJ}(t)^2 + w.F_{IJ}(t)^2\} \tag{12}$$

If we write U_{PJ} and U_{IJ} for the ultimates projected from the latest paid or incurred separately (that is $U_{PJ} = P_{J0} / F_{PJ}(t)$ and $U_{IJ} = I_{J0} / F_{IJ}(t)$), then we have:

$$U_J = \{U_{PJ}.F_{PJ}(t)^2 + w.U_{IJ}.F_{IJ}(t)^2\} / \{F_{PJ}(t)^2 + w.F_{IJ}(t)^2\} \tag{13}$$

This shows that U_J is a weighted average of U_{PJ} and U_{IJ} with weights $F_{PJ}(t)^2$ and $w.F_{IJ}(t)^2$.

For example, suppose paid and incurred amounts for the latest origin year are P_{J0} = \$200 and I_{J0} = \$1000, and suppose the fitted development curves imply that these figures are respectively 20% and 110% of ultimate, that is: $F_{PJ}(t) = 0.2$ and $F_{IJ}(t) = 1.1$. Then projecting paid and incurred separately to ultimate gives the estimates: U_{PJ} = \$200/0.2 = \$1000 and U_{IJ} = \$1000 / 1.1 = \$909. Suppose further that the analysis described in Section 2.3.3 indicates that incurred sums of squares should receive a weight of 0.2 relative to paid (that is, $w = 0.2$), then we have : $F_{PJ}(t)^2 = 0.04$ and $w.F_{IJ}(t)^2 = 0.24$. Equation 13 then gives a combined estimated ultimate: U_J = (0.04 * \$1000 + 0.24 * \$909) / 0.28 = \$922.

Note that in practice, the model can be set up in Excel, and least-squares estimation carried out using the Excel solver. It is not necessary to evaluate the formulas given above: the estimated ultimates are parameters of the model that are found by the Excel solver.

2.4.2 Model for ultimate in terms of premium or other exposure information

From Equation 12, it is clear that the estimated ultimate for the final origin year will be sensitive to the values of the two observations P_{J0} and I_{J0}. The sensitivity of the estimate to these values can be reduced (hence the reliability of the estimated ultimate increased) if total premium or some other measure of exposure can be obtained. A measure of exposure other than premium is more valuable than premium, because to make use of premium, we also have to use the estimated premium rate index for the latest year (Q_J) and this is already used in $F_{PJ}(t)$ and $F_{IJ}(t)$ (through equations such as 6 and 7). Premium and exposure data will clearly also be useful for other origin years, but it is for the latest few origin years that this additional information is most valuable.

First we consider using a measure of exposure other than premium. This might be, for example, gross tonnage in a marine account, or total payroll in workers compensation. The exposure for origin year j is denoted X_j. If there are no cycles or trends in the ultimate loss per unit of exposure, then we have:

$$U_j = r.X_j + \text{random error.} \tag{14}$$

Here, the parameter r represents the mean ultimate loss per unit of exposure.

However, it could be that there is a trend in ultimate loss per unit of exposure. We should at least expect an inflationary trend if the exposure measure is not in dollars. In this case, we could try a model of the form $r_j = \exp(\varrho_0 + \varrho_1.j)$ (where ϱ_0 and ϱ_1 are parameters assumed to take constant values across all origin years). We can also allow for the possibility that the ultimate loss per unit of exposure varies with the underwriting cycle:

$$r_j = \exp(\varrho_0 + \varrho_1.j + \varrho_2.Q_j). \tag{15}$$

Using this model, Equation 14 becomes:

$$U_j = X_j.\exp(\varrho_0 + \varrho_1.j + \varrho_2.Q_j) + \text{random error.} \tag{16}$$

In the event that the only measure of exposure available is premium (denoted Prem_j) then underwriting cycle effects need to be removed from this by using $X_j = \text{Prem}_j / Q_j$ in the above. The possibility that the premium rate index understates the true amplitude of the underwriting cycle is accommodated (approximately) by the inclusion of Q_j in the exponential factor: in this case the parameter ϱ_2 will be lower than it would be if Q_j correctly reflected the amplitude.

Another possibility is to remove the exponentiation from Equation 16 (so any inflationary trend is approximated as linear) to give:

$$U_j = \text{Prem}_j.(\varrho_2 + \varrho_1.j/Q_j + \varrho_0/Q_j) + \text{random error.} \tag{17}$$

If premium takes several years to develop to ultimate (as is often the case in London market business because of profit-sharing, reinstatement premiums, retrospective experience rating, end-of-term exposure adjustments, etc.), then Prem_j could be obtained by applying a simple projection method (such as chain ladder) to the premium development array.

In all the above equations the "random error" term reflects real variation in loss experience from one origin year to another. As a first approximation, it is probably reasonable most of the time to assume that the variance of this is proportional to the expected ultimate U_j and to approximate this as being proportional to Prem_j/Q_j. Proportionality of variance to expected value implies that the coefficient of variation is inversely proportional to the square-root of the expected ultimate, reflecting the diversification benefit of large portfolios.

2.4.3 Use of exposure information in curve-fitting

To make use of the premium (or other exposure) information in curve-fitting by least-squares, we need to add a further term to the sum of squares that is minimized (see Section 2.3.3). From Equation 17, and the above assumption on the approximate variance of the random error term, the additional sum of squares in respect of exposure is given by:

$$\text{RSS}_X = \text{Prem}_0.\Sigma_j \{U_j - \text{Prem}_j.(\varrho_2 + \varrho_1.j/Q_j + \varrho_0/Q_j)\}^2 .Q_j / \text{Prem}_j. \tag{18}$$

Here, summation is over all origin years j. Prem_0 represents the mean value of Prem_j across all origin years (or some other suitable value, e.g., $\text{Prem}_0 = \text{Prem}_J$). This factor is included to offset the factor Q_j/Prem_j applied to each term in the sum: it ensures that RSS_X is in units of "dollars-squared" and of the same order of magnitude as RSS_P and RSS_I. The total sum of squares to be minimized becomes $\text{RSS}_P + w_I.\text{RSS}_I + w_X.\text{RSS}_X$, where RSS_P and RSS_I are given by Equations 8 and 9, and appropriate values for the weights w_I and w_X can be determined iteratively using the principles described in Section 2.3.3.

For example, consider again the case of annual paid and incurred run-off arrays for 10 origin years, so that $n_P = n_I = 55$. We now have ten additional pieces of information: the estimated ultimate premiums Prem_j, denoted by $n_X = 10$. Once again using the Weibull model described in Section 2.3.2: the paid development curves depend on 13 parameters (c_P, σ_0, σ_1 and U_1, $U_2 \ldots U_{10}$), the incurred development curves depend on 17 parameters (the same 13 as in the paid model plus s_R, c_R, β_0 and β_1) and the exposure model depends on 13 parameters (the ten ultimates and the three rho-parameters of Equation 17). Parameters that feature in more than one of the three components can be counted in proportion to the number of data-points in each component. For example, as the parameter U_1 is determined using all 120 data-points, it is counted as 55/120 of a parameter in the paid model, 55/120 of a parameter in the incurred model, and 10/120 of a parameter in the exposure model. On this basis we have: $p_P = 10 * 55/120 + 3 * 55/110 = 6.08$, $p_I = 10 * 55/120 + 3 * 55/110 + 4 = 10.08$, and $p_X = 10 * 10 / 120 + 3 = 3.83$, which gives the correct total number of distinct parameters: $p_P + p_I + p_X = 20$. (A more refined method of counting parameters is proposed and discussed in Section 2.5.1.)

On the first iteration, we estimate the parameters by minimizing $\text{RSS}_P+\text{RSS}_I+\text{RSS}_X$ (that is, using equal weights initially). Suppose this produces parameter values that give $\text{RSS}_P = 102$ and $\text{RSS}_I = 308$ and $\text{RSS}_X = 50$, so the total minimized residual sum of squares is 460. (Note that RSS_P and RSS_I are necessarily higher than they were when RSS_X was not considered.) Initial estimates of the variances are then: $\sigma^2_P = 102/(55–6.08) = 2.09$ and $\sigma^2_I = 308/(55–10.08) = 6.86$ and $\sigma^2_x = 50 / (10 – 3.83) = 8.1$.

For the second iteration, we minimize $(\text{RSS}_P/2.09)+(\text{RSS}_I/6.86)+(\text{RSS}_X/8.1)$, which is equivalent to minimizing: $\text{RSS}_P + 0.304 * \text{RSS}_I + 0.258 * \text{RSS}_X$. The values of RSS_P, RSS_I and RSS_X given by minimizing this weighted sum are then used to calculate revised estimates of σ^2_P, σ^2_I and σ^2_X, and the weights for the third iteration calculated using $w_I = \sigma^2_P/\sigma^2_I$ and $w_X = \sigma^2_P/\sigma^2_X$. Convergence usually occurs after a few iterations.

2.4.4 Effect of exposure information on projected ultimate for latest origin year

In Section 2.4.1 we considered the estimated ultimate for the latest origin year (U_J) obtained using just two pieces of information for that origin year: P_J and I_J. We now consider the estimate of U_J obtained by, in addition, using the premium data as described in 2.4.3.

This latest origin year now contributes three terms to the weighted sum of squares:

$$\{P_{J0} - U_J.F_{PJ}(t)\}^2 + w_I.\{I_{J0} - U_J.F_{IJ}(t)\}^2 + w_X.\{U_J - \text{Prem}_J.R_J\}^2\ .Q_J.\text{Prem}_0/\text{Prem}_J \qquad (19)$$

(Here R_J denotes $(\varrho_2 + \varrho_1.\text{J}/Q_J + \varrho_0/Q_J)$, which can be regarded as the expected ultimate loss ratio for the latest origin year.)

Since U_J does not appear in any other terms of the total sum of squares, it can be adjusted to minimize the sum of the above three terms. It is easily proved that this gives:

$$U_J = \{P_{J0}.F_{PJ}(t) + w_I.I_{J0}.F_{IJ}(t) + w_X.R_J.Q_J.\text{Prem}_0\} / \{F_{PJ}(t)^2 + w_I.F_{IJ}(t)^2 + w_X.Q_J.\text{Prem}_0/\text{Prem}_J\}. \qquad (20)$$

If we write U_{PJ}, U_{IJ} and U_{XJ} for ultimates estimated respectively from paid, incurred and premium data separately (that is $U_{PJ} = P_{J0} / F_{PJ}(t)$, $U_{IJ} = I_{J0} / F_{IJ}(t)$ and $U_{XJ} = \text{Prem}_J.R_J$), then we have:

$$U_J = \{U_{PJ}.F_{PJ}(t)^2 + w_I.U_{IJ}.F_{IJ}(t)^2 + w_X.U_{XJ}.Q_J.\text{Prem}_0/\text{Prem}_J\} / \{F_{PJ}(t)^2 + w_I.F_{IJ}(t)^2 + w_X.Q_J.\text{Prem}_0/\text{Prem}_J\}. \qquad (21)$$

This shows that U_J is now a weighted average of $U_{PJ,}$ U_{IJ} and U_{xJ}.

For example, suppose paid and incurred amounts for the latest origin year are P_{J0} = \$200 and I_{J0} = \$1000, and suppose the fitted development curves imply that these figures are respectively 20% and 110% of ultimate, that is: $F_{PJ}(t) = 0.2$ and $F_{IJ}(t) = 1.1$. Then projecting paid and incurred separately to ultimate gives the estimates: U_{PJ} = \$200/0.2 = \$1000 and U_{IJ} = \$1000 / 1.1 = \$909. In addition, suppose we have Prem_J = \$1100 and, from the fitted loss-ratio model, R_J = 104%, so that U_{XJ} = \$1144. Suppose further that the analysis described in the previous section converges to w_I = 0.29 and w_X = 0.24. If the latest origin year is believed to be at the mid-point of the underwriting cycle (so $Q_J = 1$) and we have used the normalizing factor $\text{Prem}_0 = \text{Prem}_J$, then we have: $F_{PJ}(t)^2$ = 0.04, $w_I.F_{IJ}(t)^2 = 0.35$ and $w_X.Q_J.\text{Prem}_0/\text{Prem}_J = 0.24$. The above formula then gives a final estimate: U_J = (0.04 * \$1000 + 0.35 * \$909 + 0.24 * \$1144) / 0.63 = \$1004.

For earlier origin years, the influence of the exposure information will be lower because the number of terms in the sum of squares relating to paid and incurred development data is higher for earlier years, while the number of terms relating to exposure data remains at one for each origin year.

2.5 Parameter Counts and Significance Tests

2.5.1 Parameter counts

In the examples of Sections 2.3.3 and 2.4.3, the total parameter count was apportioned between the different sub-sets of data (paid, incurred and exposure data) in proportion to the number of data-points. The example of Section 2.3.3 has an equal volume of paid and incurred data (55 paid and 55 incurred observations) and no exposure data, so each parameter that features in both paid and incurred models was counted as half a parameter in each of these models. The example of Section 2.4.3 has 10 exposure observations in addition to the 55 paid and 55 incurred observations. Each parameter that contributes to all three parts of the model was counted as 10/120 of a parameter in the exposure part, and 55/120 in each of the paid and incurred parts.

The rationale for splitting parameter counts in this way is that it approximately reflects the relative influence of each type of data in determining the value of the parameter. This can be further refined by taking account of the relative weights used in the total weighted sum of squares that is minimized. For example, if we have just 55 paid and 55 incurred observations (no exposure data), and the incurred data is given a weight of 0.5 relative to the paid data (that is, $w_I = 0.5$), then each incurred observation has (on average) only half the influence of each paid observation in determining parameter values. On this basis, a parameter whose value is determined from both paid and incurred data would be counted as 2/3 determined from paid data and 1/3 from incurred data.

In general using this method, each parameter contributes to p_P, p_I, and p_X in proportion to n_P, $w_I.n_I$, and $w_X.n_X$. Each parameter must have a total count of one, so for a parameter estimated from all three data sources, the contributions to p_P, p_I, and p_X are respectively:

$$n_P \,/\, (n_P + w_I.n_I + w_X.n_X), \quad w_I.n_I \,/\, (n_P + w_I.n_I + w_X.n_X), \quad w_X.n_X \,/ (n_P + w_I.n_I + w_X.n_X).$$

For a parameter that is not estimated from all three sources, the corresponding term(s) must be omitted from the denominator of these expressions so that the total is always one for each parameter. No rigorous theoretical justification is claimed for this method of counting parameters: it is proposed as a reasonable pragmatic approach.

The method for counting parameters used in Sections 2.3.3 and 2.4.3 is as above but with the weights w_I and w_X omitted. Including these weights in the parameter counts slightly complicates the fitting process because the relative weights (estimated from the residual sums of squares as described in Sections 2.3.3 and 2.4.3) change at each iteration, so the parameter counts will also change at each iteration.

To illustrate this, consider again the example of Section 2.4.3. For the first iteration, we use $w_I = w_X = 1$. Therefore the parameter counts are initially as in Section 2.4.3: $p_P = 6.08$, $p_I = 10.08$, $p_X = 3.83$ (giving a total parameter count of 20). Estimates of variances obtained from the first iteration are then (exactly as in Section 2.4.3): $\sigma^2_P = 102/(55–6.08) = 2.09$ and $\sigma^2_I = 308/(55–10.08) = 6.86$ and $\sigma^2_x = 50/(10–3.83) = 8.1$.

For the second iteration, the relative weights become $w_I = 2.09 / 6.86 = 0.304$, and $w_X = 2.09 / 8.1 = 0.258$, so we minimize $RSS_P + 0.304 * RSS_I + 0.258 * RSS_X$ (which again, is exactly as in Section 2.4.3). However, unlike in Section 2.4.3, the parameter counts used in estimating variances from the results of the second iteration should now factor in the weights used in the second iteration. Parameter counts that factor in these weights are:

$p_P = 3 * \{55 / (55 + 0.304 * 55)\} + 10 * \{55 / (55 + 0.304 * 55 + 0.258 * 10)\}$,

$p_I = 0.304 * p_P + 4$,

$p_X = 10 * \{0.258 * 10 / (55 + 0.304 * 55 + 0.258 * 10)\} + 3$.

These evaluate to $p_P = 9.70$, $p_I = 6.95$, $p_X = 3.35$. (As a check, we see that the total of these parameter counts is still 20.) Estimates of variances from the second iteration are then: $\sigma^2_P = RSS_P/(55 – 9.70)$, $\sigma^2_I = RSS_I/(55 – 6.95)$ and $\sigma^2_x = RSS_X / (10 – 3.35)$. These give weights w_I and w_X for the third iteration, and hence parameter counts used in estimating variances from the results of the third iteration.

2.5.2 Statistical significance tests

When an additional parameter is introduced into a model, the minimized residual sum of squares is inevitably smaller than it was before the new parameter was introduced. (This is because the model without the additional parameter is equivalent to a model in which the additional parameter is set to zero. When the new parameter is introduced, it is no longer constrained to take the value zero. When non-zero values are allowed in minimizing the residual sum of squares, it is extremely unlikely that the minimum will occur at exactly the value zero.) So by introducing an increasing number of parameters, the quality of fit (as measured by the residual sum of squares) can be progressively improved until the number of parameters is equal to the number of observations: when this occurs a perfect fit is possible and the residual sum of squares becomes zero.

Clearly we should try to avoid over-fitting: that is, we should try to avoid including parameters that reflect only the random variation of the particular dataset rather than genuine underlying effects. The purpose of statistical significance tests is to avoid over-fitting: parameters are included in a model only if they are statistically significant.

In least squares estimation, an appropriate significance test is based on the size of the decrease in the minimized residual sum of squares (RSS) when a new parameter is introduced. If the minimized RSS reduces only slightly, then the new parameter may not be statistically significant. To judge whether a decrease in the minimized RSS is statistically significant, it should be compared to the mean RSS per "degree of freedom."

To illustrate, consider again the example of Section 2.3.3 based on 55 paid observations and 55 incurred observations. Suppose that initially we fit a model with 17 parameters and that we give equal weight to incurred and paid data so parameters are estimated by minimizing $RSS_P + RSS_I$. Suppose (as in Section 2.3.3) that the minimized value is 400.0. Now suppose that an additional parameter is introduced into the model (this might, for example, be the parameter β_I of Equation 6), and that the minimized RSS with this new parameter included is 397.0. Is the new parameter statistically significant?

To answer this, we note that the new parameter caused a decrease of 3.0 in the minimized RSS. If this is judged to be a large decrease, then we conclude that it is unlikely to have been caused purely by chance and therefore that the new parameter is statistically significant. To judge whether the decrease of 3.0 is large (statistically significant) or small (insignificant) we need to compare it to something else. At first sight, it might seem that a suitable quantity to compare this decrease to is the mean-squared residual. There are 110 residuals in total (equal to the number of data-points) so the mean-squared residual is 397/110 (= 3.61) when the new parameter is included, and 400/110 (=3.64) when it is not included. Clearly the mean-squared residual decreases as the number of parameters increases. To allow for this, the denominator used in calculating the mean needs to be adjusted for the number of parameters in the model. Instead of dividing by the number of observations, we should divide by the number of "degrees of freedom", which is defined as the number of observations less the number of fitted parameters. The number of parameters was 17 before then new parameter was introduced and 18 afterwards, so the numbers of degrees of freedom are respectively 93 and 92. The mean-squared residual per degree of freedom is therefore 400/93 (= 4.30) before the new parameter is introduced and 397/92 (=4.32) after. To judge whether the new parameter is statistically significant, we should compare the change in the RSS of 3.0 to the mean value 4.32. Since the change in the RSS is less than the mean RSS per degree of freedom, the new parameter is not statistically significant.

Note that the change in the RSS is compared to the mean obtained from the more general model (4.32 in this example) not the mean obtained from the model excluding the parameter (4.30 in this example). This is because, if a parameter is statistically significant, the mean RSS from the model with the parameter excluded would be wrongly inflated by the exclusion.

If the decrease in the RSS obtained by introducing an additional parameter is less than the mean RSS per degree of freedom after the parameter has been introduced (as in the above example), then the additional parameter is not statistically significant. However, a change in the RSS that exceeds the mean RSS per degree of freedom is not necessarily conclusive evidence of statistical significance. Clearly, the greater the ratio of change in RSS to RSS per degree of freedom, the greater the statistical significance of the new parameter. (In the above example, this ratio is 3.0/4.32 = 0.69.) This ratio is known as the "F-ratio" and, to help judge its statistical significance, it can be compared to a theoretical F-distribution. In idealized circumstances, the theoretical F-distribution is the probability distribution of an F-ratio under the hypothesis that the additional parameter is equal to zero. Although this is not exactly the case in practice, the theoretical F-distribution remains a useful tool in judging the statistical significance of F-ratios. If an F-ratio is in the extreme right tail of the theoretical F-distribution, this is evidence against the hypothesis that the true value of the parameter is zero. In other words, the parameter is statistically significant. In our example, the appropriate theoretical F-distribution is that with 1 and 92 degrees of freedom (1 because one additional parameter has been introduced, 92 because after introducing the additional parameter, the RSS has 92 degrees of freedom, that is, 110 - 18). The 95th percentile of the theoretical F distribution with 1 and 92 degrees of freedom is 3.95. This means that if the true value of the new parameter is zero, there is only a 5% chance that the F-ratio would be as high as 3.95. So an F-ratio in excess of 3.95 is strong evidence that the true value of the parameter is non-zero. An F-ratio above the 95th percentile is usually judged to be statistically significant. A value above the 90th percentile (2.77 in this example) would also usually be regarded as statistically significant, but with a lower degree of confidence. Any F-ratio greater than 1 provides some evidence that the parameter is in fact non-zero, but clearly the lower the F-ratio, the weaker the evidence.

In carrying out F-tests on weighted sums of squares, it is important to ensure that the weights are the same in both the numerator and denominator of the F-ratio. It would be wrong to compare a change in $\text{RSS}_P + w_I.\text{RSS}_I$ to a mean value of $\text{RSS}_P + w_I'.\text{RSS}_I$ unless w_I' is equal to w_I. If the weights are not equal, then we are not comparing like with like and the F-ratio is meaningless.

In the present paper, a model is fitted iteratively (with the number of parameters fixed) until the relative weight w_I converges (as described in 2.3.3). Additional parameters should then be introduced, initially with no change in w_I in order to carry out a valid F-test. If the F-test shows the additional parameters to be statistically significant, further iterations can then be carried out with the additional parameters included until w_I converges to a new value.

2.5.3 Extra parameters for first few development years

When fitting theoretical run-off curves to discrete aggregate development data, it often happens that the fit is relatively poor for the first one or two development years. This occurs because the first development year contains a mixture of actual delays from the accident date to the end of the development year (depending on the distribution of accident occurrence dates in the accident year). In the case of underwriting year cohorts, the situation is further complicated by the range of possible policy inception dates within the underwriting year. These effects are approximately taken into account by using the average values of t given in Table 2 (Section 2.3.3). These approximations are often poor for the first one or two development years. The accuracy of these approximations increases in later development years because the variation in accident dates within the first year is proportionately a smaller part of the total delay.

This phenomenon was observed in Wright (1989) where it was accommodated by introducing some additional parameters for the first few development periods. The same refinement can easily be introduced in the present model. It is usually only the first two development years that are significantly affected. To allow for these, up to four additional parameters are required: two for paid development and two for incurred development.

In the model of Section 2.3.2 (Equations 8 and 9) the paid and incurred expected values are modeled respectively as $U_j.F_{Pj}(t)$ and $U_j.F_{Ij}(t)$. Here, j is the origin year, t is the average development delay (given by the approximations in Table 2), $F_{Pj}(t)$ and $F_{Ij}(t)$ are the run-off curves given by Equations 4, 5, 6, and 7, and U_j is the ultimate for origin year j.

We now introduce additional parameters θ_{P0}, θ_{P1}, θ_{I0}, θ_{I1}. The subscripts 0 and 1 indicate that these parameters apply to development years 0 and 1. For these two development years, the expected cumulative paid values are modeled as $\exp(\theta_{P0}).U_j.F_{Pj}(t)$ (where t = 0.5 for accident years, 0.333 for underwriting years) and $\exp(\theta_{P1}).U_j.F_{Pj}(t)$ (where t = 1.5 for accident years, 1.0 for underwriting years), and similarly for cumulative incurred. Note that with this form of model:

- Each θ-parameter may take any real value (positive or negative).
- The value zero for a θ-parameter corresponds to the case where no adjustment is necessary (the factor exp(θ) is then one so can be omitted).
- The θ-parameters can be determined in the same was as any other parameter of the model: by least squares estimation.
- Statistical hypothesis tests can be carried out (as described in 2.5.2) to determine whether or not these additional parameters are necessary.

3. EXAMPLE ANALYSIS

3.1 Data

To illustrate the methods described in Section 2, they are applied to the development data given in Appendix A. This is based on actual data, covering underwriting years 1993 to 2006. The class of business and other details are not given here to preserve confidentiality. The numbers of paid and incurred data-points are $n_P = n_I = 105$, giving a total of 210. The premium rate index (given in Section A.1.4. of the appendix) is an estimate obtained by applying conventional projection methods to the triangles to find estimated ultimate premiums and claims. The premium rate index was then calculated as the ratio of estimated ultimate premiums to estimated ultimate claims. This was adjusted by a constant factor so the mean value of the index Q_j over the 14 underwriting years is one. (The reliability of this method of calculating a premium rate index is discussed in Section 4.2.)

3.2 Weibull model

3.2.1 Constant development pattern

First we fit the Weibull model with all parameters fixed at constant values across all origin years. In other words, we assume initially that the run-off pattern is the same for all origin years with no dependence on the underwriting cycle. This model has a total of 19 parameters: s_P, c_P, s_R, c_R, b, and $U_1 \ldots U_{14}$. First estimates of the parameters are given by minimizing the un-weighted total residual sum of squares $\text{RSS}_P+\text{RSS}_I$. (This is Equation 9 in the case $w_I = 1$, where RSS_P and RSS_I are given by Equation 8.) Allocating the 19 parameters between paid and incurred data as described in Section

2.5.1 gives $p_P = 8$, $p_I = 11$.

Results are shown in the second column (iteration number 1) of Table 3. The residual sums of squares are in millions. Initial estimates of typical paid and incurred variances from Equation 10 are: $\sigma^2_P = 276.7/97 = 2.85$ and $\sigma^2_I = 351.3/94 = 3.74$. The fact that σ^2_I is higher than σ^2_P indicates that the Weibull curve does not fit as closely to the incurred data as to the paid data. This gives a weight for the second iteration of $w_I = 2.85/3.74 = 0.763$. Parameter counts (using the method described in 2.5.1) are then:

$p_P = 16 * 105 / (105 + 0.763 * 105) = 9.07$ and $p_I = 3 + 0.763 * p_P = 9.93$.

Results of minimizing the weighted residual sum of squares are given in the corresponding column (iteration number 2) of Table 3. These results give new estimates: $\sigma^2_P = 263.4/95.93 = 2.75$, $\sigma^2_I = 366.6/95.07 = 3.86$ hence $w_I = 2.75 / 3.86 = 0.712$ for the third iteration. Continuing in this way, convergence occurs in five iterations. Using the formula for the mean of a Weibull distribution (see Section 2.2.2), the final values of the Weibull parameters imply a mean reporting delay of 1.8 years and a mean payment delay of 2.7 years. The final column of Table 3 shows the ratio of the ultimates from the converged Weibull model to the basic chain ladder ultimates obtained from just the incurred data.

Table 3: Weibull curves with constant parameters

Iteration	1	2	3	4	5	U_j as % of ICL
w_I	1	0.763	0.712	0.699	0.700	
p_P	8	9.07	9.35	9.42	9.41	
p_I	11	9.93	9.65	9.58	9.59	
RSS_P	276.7	263.4	260.2	260.2	260.2	
RSS_I	351.3	366.6	370.9	370.9	370.9	
$RSS_P+w_I.RSS_I$	628.0	543.2	524.3	519.6	520.0	
s_P	3.00	3.01	3.02	3.02	3.02	
c_P	1.41	1.40	1.40	1.40	1.40	
s_R	2.01	1.93	1.91	1.91	1.91	
c_R	1.24	1.24	1.24	1.24	1.24	
b	0.98	0.89	0.87	0.87	0.87	
U_1	13,913	13,927	13,930	13,930	13,930	102.7%
U_2	19,130	19,165	19,174	19,174	19,174	101.2%
U_3	11,200	11,217	11,221	11,221	11,221	99.6%
U_4	10,995	10,980	10,976	10,976	10,976	101.8%
U_5	12,982	12,960	12,954	12,954	12,954	93.8%
U_6	26,159	25,838	25,755	25,755	25,755	89.2%
U_7	68,255	68,256	68,257	68,257	68,257	94.8%
U_8	142,745	143,104	143,197	143,197	143,197	98.1%
U_9	128,173	128,703	128,841	128,841	128,841	101.6%
U_{10}	65,742	65,949	66,003	66,003	66,003	92.4%
U_{11}	4,445	4,377	4,359	4,359	4,359	92.0%
U_{12}	4,440	4,458	4,463	4,463	4,463	103.1%
U_{13}	7,422	7,180	7,112	7,112	7,112	86.5%
U_{14}	24,784	24,252	24,093	24,093	24,093	81.2%
ΣUj	540,385	540,365	540,334	540,334	540,334	96.4%

3.2.2 Varying bias factor

Next some of the parameters are allowed to vary with the underwriting cycle using the model of Equations 6 and 7. First we allow just the parameter b to vary, so instead of a single parameter b, we now have two parameters β_0 and β_1 (see Equation 6). The additional parameter relates to incurred data only so p_I increases by 1 giving: $p_P = 9.41$, $p_I = 10.59$.

Table 4: Weibull curves with varying *b*-parameter

Iteration	1	2
w_I	0.700	0.697
p_P	9.41	9.43
p_I	10.59	10.57
RSS_P	260.3	260.3
RSS_I	368.7	368.7
$RSS_P + w_I.RSS_I$	518.5	517.4
s_P	3.02	3.02
c_P	1.40	1.40
s_R	1.90	1.90
c_R	1.24	1.24
β_0	-0.053	-0.053
β_1	0.191	0.191
ΣU_j	537,411	537,411

The second column (iteration 1) of Table 4 shows least squares results obtained using the same value w_I = 0.700 as used in the model with b constant. The additional parameter β_1 causes the weighted RSS to fall from 520.0 to 518.5. An approximate test of statistical significance of the additional parameter is the F-test described in Section 2.5.2. This is based on the ratio of the decrease in the weighted RSS per additional parameter (which is 1.5 in this case, as there is only one additional parameter) to the mean RSS per degree of freedom in the model with 20 parameters, which is 518.5 / (210 – 20) = 2.7. If the additional parameter (β_1) is actually zero, then this ratio has approximately an F-distribution with 1 and 190 degrees of freedom, so a value in the extreme right-tail of the F-distribution would be evidence against the hypothesis that β_1 is zero. In this case, the ratio (1.5/2.7) is less than 0.5, which is not an extreme value compared to an F-distribution, so the parameter β_1 is not statistically significant. Nevertheless, β_1 being positive (0.191) is weak evidence that case estimates are strengthened in harder markets.

3.2.2 Varying paid development time-scale parameter

Since the F-test indicates no strong evidence that b varies with the underwriting cycle, we next try a model in which b is constant, but the paid-development scale parameter (s_P) is allowed to vary as in Equation 7. We start (iteration 1) with w_I = 0.700 as in Table 3, so the additional parameter contributes 1/1.700 to p_P and 0.700/1.700 to p_I, to give p_P = 10.00, p_I = 10.00. The weighted RSS becomes 510.3, which is a decrease of 9.7. This gives the *F*-ratio 9.7/(510.3/190) = 3.6, which is high, indicating that the additional parameter is statistically significant this time. σ_1 being less than zero (-0.199) indicates that s_{Pj} decreases as the premium index Q_j increases, that is, payment delays tend to be shorter in harder markets. Three further iterations are necessary for convergence as shown in Table 5.

Table 5: Weibull curves with varying payment delay parameter

Iteration	1	2	3	4
w_I	0.700	0.665	0.656	0.653
p_P	10.00	10.21	10.27	10.28
p_I	10.00	9.79	9.73	9.72
RSS_P	248.5	246.2	245.6	245.6
RSS_I	373.8	377.2	378.1	378.1
$RSS_P+w_I.RSS_I$	510.3	497.0	493.5	492.6
σ_0	1.015	1.015	1.015	1.015
σ_1	-0.199	-0.201	-0.201	-0.201
c_P	1.40	1.39	1.39	1.39
s_R	1.63	1.63	1.62	1.62
c_R	1.26	1.26	1.26	1.26
b	0.66	0.65	0.65	0.65

Next we try a model in which both b and s_P are allowed to vary (as in Equations 6 and 7). Although it seemed that β_1 was not statistically significant when s_P was held constant, it is possible that when both b and s_P are allowed to vary, both are statistically significant. With the weight w_I fixed at 0.653, the additional parameter (β_1) causes a decrease in the weighted RSS from 492.6 to 485.7, a decrease of 6.9. The *F*-ratio is 6.9 / (485.7/189) = 2.7, which is close to the 89th percentile of the corresponding *F*-distribution. Since this is not an extremely high percentile, the statistical significance of the parameter is not clear. In this case, we continue with this model for now, and re-test the significance of the β_1 parameter at a later stage. Note that the final column of Table 6 (ratio of cyclical ultimates to incurred CL ultimates) now shows a fairly clear cyclical pattern.

Table 6: Weibull curves with varying payment delay and *b* parameters

Iteration	1	2	U_j as % of incurred CL
w_I	0.653	0.655	
p_P		10.27	
p_I		10.73	
RSS_P	243.8	243.9	
RSS_I	370.4	370.2	
$RSS_P+w_I.RSS_I$	485.7	486.4	
σ_0	0.986	0.986	
σ_1	-0.251	-0.251	
c_P	1.40	1.40	
s_R	1.70	1.70	
c_R	1.25	1.25	
β_0	-0.098	-0.097	
β_1	0.513	0.513	
U_1	13,598	13,598	100.2%
U_2	19,057	19,056	100.6%
U_3	11,056	11,056	98.2%
U_4	10,792	10,793	100.1%
U_5	12,804	12,804	92.7%
U_6	25,661	25,664	88.9%
U_7	68,454	68,455	95.1%
U_8	145,666	145,662	99.8%
U_9	128,080	128,076	101.0%
U_{10}	63,611	63,610	89.1%
U_{11}	3,343	3,344	70.6%
U_{12}	3,223	3,224	74.5%
U_{13}	5,815	5,817	70.7%
U_{14}	18,283	18,284	61.6%
ΣU_j	529,443	529,444	94.5%

Table 7 shows the implied variation of s_P and b across the origin years: these values are calculated from Equations 6 and 7 using the parameter estimates given in Table 6. Comparing the hard market of 2003-2004 to the soft market of 1998-2001 these results imply:

- Approximately a one-third reduction in payment delays in the hard market (s_P decreases from about three in the soft market to about two in the hard market).
- More than doubling of case estimates in the hard market (b increased from 0.7 in the soft

market to about 1.7 in the hard market conditions). However, the statistical significance of the β_1 parameter was questionable so that this apparent cyclical effect might in fact be caused by random variation: results presented in the next sub-section suggest that this is in fact the case.

Table 7: Weibull curves with varying payment delay and *b* parameters

Origin year (j)	Q_j - 1	s_P	b
1993	0.17	2.57	0.99
1994	-0.32	2.90	0.77
1995	-0.18	2.81	0.83
1996	-0.19	2.81	0.82
1997	-0.32	2.90	0.77
1998	-0.47	3.01	0.71
1999	-0.49	3.03	0.70
2000	-0.60	3.11	0.67
2001	-0.42	2.98	0.73
2002	-0.25	2.86	0.80
2003	1.14	2.02	1.63
2004	1.34	1.92	1.81
2005	0.18	2.56	1.00
2006	0.41	2.42	1.12

3.2.2 Use of Premium data

We could now test whether there is any evidence that reporting delays also vary with the cycle by using a model like Equation 7, but for the scale parameter of reporting delay s_R. However, before doing this, we test the effect of using premium data as described in Section 2.4.3. The number of ultimate premium data-points is 14 so the total number of data-points increases to 224. First we try just one additional parameter, ϱ_0 (that is, we use Equation 17 with parameters ϱ_1 and ϱ_2 set to zero). Convergence occurs in four iterations as shown in Table 8. The value 1.066 (or 106.6%) for ϱ_0 represents the mean ultimate loss ratio on the assumption that, after on-leveling premiums using the premium rate index Q_j, the mean ultimate loss ratio is the same for all origin years.

Table 8: Weibull curves with varying payment delay and *b* parameters, using premium data

Iteration	1	2	3	4
w_I	0.655	0.701	0.726	0.737
w_X	1	1.606	1.884	2.005
	9.64	9.08	8.82	8.71
	10.32	10.36	10.40	10.42
	2.04	2.57	2.78	2.87
RSS_P	255.6	263.0	266.4	266.4
RSS_I	362.3	357.5	355.6	355.6
RSS_X	19.9	16.6	15.5	15.5
$RSS_P+w_I.RSS_I + w_X.RSS_X$	512.9	540.2	553.7	559.6
σ_0	1.029	1.038	1.041	1.041
σ_1	-0.155	-0.132	-0.125	-0.124
c_P	1.40	1.41	1.41	1.41
s_R	1.85	1.90	1.92	1.92
c_R	1.24	1.24	1.24	1.24
β_0	-0.048	-0.016	0.000	0.000
β_1	0.310	0.261	0.246	0.246
ϱ_0	1.065	1.066	1.066	1.066
ΣU_j	535,711	535,997	536,064	536,064

Since the statistical significance of the β_1 parameter was unclear when the model was calibrated using just the paid and incurred claims data, we next test the significance of this parameter when the premium data is also used in calibration. If the β_1 parameter is set to zero and least squares estimation carried out using the same weights as above (w_I = 0.737, w_X = 2.005), the effect is to increase the minimized RSS from 559.6 to 561.3. This increase is not statistically significant (F-ratio = 1.7 / (559.6 / (224 -22)) = 0.61), implying there is no clear evidence that β_1 is non-zero. Table 9 shows results for the model in which payment scale parameter s_P varies with the underwriting cycle, but the parameter b is the same across all underwriting years.

Including parameters ϱ_1 and ϱ_2 (with weights w_I = 0.751 and w_X = 2.294 as in Table 9) causes the weighted RSS to reduce from 570.4 to 569.7, which is clearly not a statistically significant reduction. Including a parameter that allows the reporting delay to vary with the underwriting cycle reduces the weighted RSS from 570.4 to 568.7, which again is not statistically significant (F-ratio = (570.4 – 568.7) / (568.7 / 202) = 0.60).

Table 9: Weibull curves with varying payment delay parameter only, using premium data

Iteration	1	2	3	4	U_j as % of ICL
w_I	0.737	0.743	0.749	0.751	
w_X	2.005	2.199	2.268	2.295	
p_P	8.71	8.60	8.54	8.52	
p_I	9.42	9.39	9.40	9.40	
p_X	2.87	3.02	3.06	3.08	
RSS_P	268.6	270.3	271.0	271.3	
RSS_I	358.8	358.0	357.6	357.4	
RSS_X	14.1	13.6	13.4	13.3	
$RSS_P+w_I.RSS_I + w_x.RSS_x$	561.3	566.2	569.2	570.4	
σ_0	1.054	1.056	1.056	1.056	
σ_1	-0.106	-0.102	-0.101	-0.101	
c_P	1.40	1.40	1.40	1.40	
s_R	1.79	1.80	1.81	1.81	
c_R	1.24	1.24	1.24	1.24	
b	0.77	0.78	0.78	0.78	
ϱ_0	1.070	1.070	1.070	1.070	
U_1	13,734	13,730	13,729	13,728	101.2%
U_2	19,064	19,058	19,056	19,055	100.5%
U_3	11,128	11,126	11,125	11,125	98.8%
U_4	10,823	10,818	10,816	10,816	100.4%
U_5	13,109	13,122	13,127	13,129	95.0%
U_6	25,984	26,002	26,013	26,017	90.2%
U_7	68,561	68,568	68,572	68,573	95.3%
U_8	143,808	143,746	143,723	143,713	98.5%
U_9	128,143	128,080	128,056	128,046	101.0%
U_{10}	66,568	66,659	66,687	66,698	93.4%
U_{11}	4,513	4,520	4,523	4,524	95.5%
U_{12}	3,839	3,839	3,839	3,839	88.7%
U_{13}	7,969	7,971	7,971	7,972	96.9%
U_{14}	20,748	20,747	20,747	20,747	69.9%
ΣU_j	537,993	537,987	537,984	537,982	96.0%

The results now show a smaller (and more plausible) amount of variation in mean payment delays with the underwriting cycle: compare Table 10 to Table 7. Table 10 also shows the implied mean payment delay in years (from the formula given in Table 1). The mean reporting delay is 1.63 years (the same for all origin years).

Table 10: Weibull curves with varying payment delay parameter only, using premium data

Origin year (j)	Q_j - 1	s_P	mean (years)
1993	0.17	2.83	2.57
1994	-0.32	2.97	2.70
1995	-0.18	2.93	2.67
1996	-0.19	2.93	2.67
1997	-0.32	2.97	2.70
1998	-0.47	3.01	2.75
1999	-0.49	3.02	2.75
2000	-0.60	3.05	2.78
2001	-0.42	3.00	2.73
2002	-0.25	2.95	2.69
2003	1.14	2.56	2.34
2004	1.34	2.51	2.29
2005	0.18	2.82	2.57
2006	0.41	2.76	2.51

3.3 Burr model

Using Burr distributions for both paid and reporting delays gives higher residual sums of squares than using the Weibull model, indicating that Burr curves provide a poorer fit to this particular dataset.

3.4 Inverse Burr model

To compare the quality of fit of the Inverse Burr and Weibull models, we fit an Inverse Burr model using the same values of the weights as in Table 8: w_I = 0.737 and w_X = 2.005. The minimized weighted RSS is 538.7 (see Table 11) which is 20.9 lower than obtained using the Weibull model (Table 8). The Inverse Burr model has two additional parameters (there are two shape parameters instead of one for both reporting and payment delays), so the decrease is 10.5 for each additional parameter. Comparing this to the RSS per degree of freedom (538.0 / 200 = 2.7) the decrease appears to be statistically significant. (Note that a formal F-test is not strictly valid here because the two models are not nested.) We conclude that the Inverse Burr model fits this particular dataset better than the Weibull model. Convergence occurs in six iterations as shown in Table 11 (intermediate results are not shown for all six iterations).

Table 11: Inverse Burr curves with varying payment delay and *b* parameters, using premium data

Iteration	1	6
w_I	0.737	0.544
w_X	2.005	1.520
p_P	9.29	10.60
p_I	11.85	10.77
p_X	2.87	2.62
RSS_P	230.0	215.8
RSS_I	379.1	395.8
RSS_X	15.1	17.1
$RSS_P + w_I.RSS_I + w_X.RSS_X$	538.7	457.3
σ_0	1.406	1.405
σ_1	-0.141	-0.148
a_P	0.27	0.26
c_P	4.00	4.12
s_R	2.76	2.82
a_R	0.31	0.29
c_R	3.29	3.44
β_0	-0.035	-0.011
β_1	0.237	0.257
ϱ_0	1.063	1.056

If the β_1 parameter is set to zero and least squares estimation carried out using the same weights as above (w_I = 0.544, w_X = 1.520), the minimized RSS increases from 457.3 to 459.0. This increase is not statistically significant (*F*-ratio = 1.7 / (457.3 / 200) = 0.74), implying (as for the Weibull model) that there is no clear evidence that β_1 is non-zero. After convergence, the final parameter values imply (using the formula for the mean of an Inverse Burr distribution from Section 2.2.2) that the mean reporting delay is 1.72 years. The mean payment delay varies with the underwriting cycle as shown in Table 13.

Including parameters ϱ_1 and ϱ_2 (with w_I and w_X as in Table 12) causes the weighted RSS to reduce from 461.2 to 460.3, which is clearly not a statistically significant reduction. Including a parameter that allows the reporting delay to vary with the underwriting cycle reduces the weighted RSS from 461.2 to 459.3, which again is not statistically significant (*F*-ratio = 1.9 / (459.3 / 200) = 0.83). Including additional parameters as described in Section 2.5.3 shows that these are statistically significant for this dataset, but for reasons of space, further results are not given here.

Table 13: Inverse Burr curves with varying payment delay parameter only, using premium data

Origin year (j)	Q_j - 1	s_P	mean (years)
1993	0.17	4.03	2.51
1994	-0.32	4.29	2.67
1995	-0.18	4.22	2.62
1996	-0.19	4.22	2.63
1997	-0.32	4.29	2.67
1998	-0.47	4.37	2.72
1999	-0.49	4.39	2.73
2000	-0.60	4.45	2.77
2001	-0.42	4.35	2.70
2002	-0.25	4.26	2.65
2003	1.14	3.56	2.22
2004	1.34	3.47	2.16
2005	0.18	4.02	2.50
2006	0.41	3.91	2.43

4. CONCLUSIONS

4.1 Commonly seen cycle dependencies

4.1.1 Variation of payment delays with the underwriting cycle

The example analysis of Section 3 shows evidence of payment delays lengthening in soft market origin years. The model described in this paper has been applied to several actual datasets for different classes of business and the finding that payment delays are longer in soft markets occurs consistently. This concurs with the findings of previous research described in Section 1 of the present paper. Possible causes are listed in Section 1.1.2.

4.1.2 Variation of case reserve strength with the underwriting cycle

Although there is no clear evidence that the case estimate bias factor varies with the underwriting cycle in the example analysis of Section 3, application of the model to other datasets has in many cases shown clear evidence that case reserves are set at higher levels in origin years with higher premium rates.

4.1.3 Variation of reporting delays with the underwriting cycle

Applications of the model to other datasets have shown, in some cases, evidence that reporting delays tend to be shorter in hard markets. This is not something that has been explicitly suggested in previous research. Some possible reasons why reporting delays might be extended in soft markets are discussed here. First we should note that what we actually measure as "reporting delay" is the time between the accident date and a case reserve being created on the insurer's claim administration system. This is the sum of two main components: (a) the true reporting delay (between loss occurrence and time when the insured, or broker, reports the loss to the insurer), and (*b*) the time between the claim being reported to the insurer and the initial case reserve being created. It is possible that the second component becomes longer in soft markets. Indeed there has been a recent high profile case in the UK in which senior insurance company executives were jailed for concealing reported claims. Having acknowledged that increased delays in this second component are possible, we now focus on reasons why the true reporting delay might increase in soft markets. There are several possibilities:

- In soft markets, insureds (and/or brokers) might be aware that they got a good deal on their insurance, and be concerned that on renewal the premium is likely to increase. For this reason, they might deliberately delay reporting valid claims until after renewal negotiations have been completed. This would border on fraudulent behavior by the insureds, but nevertheless is clearly possible.
- If cover is extended by relaxing terms and conditions in a soft market, insureds might genuinely fail to realize initially that they can claim for certain types of loss.
- If periods of cover have been extended beyond the usual one year in soft markets, it is possible that this is not correctly allowed for when compiling the aggregate run-off arrays. For example, when all policies run for one year, it would be correct to assume that if the accident date does not fall in the year the policy was written, then the loss should be allocated to the following accident year. However, if this method of allocating losses to accident years is continued when some policies run for more than one year, then development delays will appear (wrongly) to be extended.

4.2 Accuracy of premium rate index

In the example analysis, we used the reciprocal of estimated ULRs instead of a premium index.

Clearly this is not ideal. The ultimate ULR varies with claims experience, not just premium rate variation. A high ULR is not necessarily indicative of a soft market: it might occur simply because of unusually high loss experience. Although the results appear to show longer development tails in softer markets, could it be that in fact all we are seeing is longer development patterns when losses are exceptionally high? The results could partly reflect this, but the fact that the estimated ULRs do broadly follow a cyclical pattern (rather than just random variation) suggests that most of the variation in ULRs reflects variation in premium rates with the underwriting cycle.

Where the models show significant cyclical variation in run-off patterns, the estimated reserve for the latest year will clearly be sensitive to the value of the premium rate index for that year, and this is the most difficult year to get an accurate fix on. A worthwhile area for future research would be to predict premium rate variation of the underwriting cycle. For example, we might expect premium rates next year to be related to measurable quantities such as the number of new insurance company start-ups this year, or the amount of new capital in the insurance industry. If the underwriting cycle can be predicted from such quantities (even if only one year ahead), then the accuracy of reserves could be improved by using these quantities directly in the reserving model instead of (or as well as) the estimated premium rate index Q_j. There is a substantial literature on the underwriting cycle and its causes: this could point to suitable alternative variables to use instead of Q_j.

5. REFERENCES

[1] Conger, R. "Actuaries: Deeper? Broader? Or Uncertain?" Presented at the 2002GIRO Convention.

[2] General Insurance Reserving Issues Taskforce (chaired by A.R. Jones), "A Change Agenda for Reserving," *British Actuarial Journal* 2006, Vol. 12, No. 3:435-599.

[3] Line, N. et al., "The Cycle Survival Kit: An Investigation into the Reserving Cycle and Other Issues," GIRO Working Party Report 2003, www.actuaries.org.uk/files/pdf/proceedings/giro2003/Line.pdf.

[4] Wright, T. S., "A Stochastic Method for Reserving in General Insurance," *Journal of the Institute of Actuaries* 1990, Vol. 117:677-731 l

APPENDIX A – DATA FOR EXAMPLE

A.1.1 Paid claims

1993	1,051	3,093	6,251	9,490	12,341	12,738	13,246	13,267	13,319	13,346	13,412	13,441	13,443	13,551
1994	1,102	6,112	10,284	14,047	15,690	16,505	17,539	18,071	18,457	18,729	18,676	18,679	18,917	
1995	824	2,467	3,938	7,311	8,538	9,537	10,849	11,016	11,078	11,111	11,117	11,132		
1996	726	2,876	4,817	7,279	8,783	9,694	9,782	10,006	10,283	10,508	10,548			
1997	443	2,293	5,113	7,381	10,016	10,684	11,102	12,453	12,868	13,084				
1998	383	5,169	12,181	16,846	17,852	19,817	20,974	21,331	22,193					
1999	3,705	17,901	27,656	40,224	51,593	61,108	62,227	63,858						
2000	6,839	30,329	55,327	86,931	112,799	127,208	133,792							
2001	5,048	29,082	56,406	83,434	101,096	111,152								
2002	2,936	10,127	22,969	42,236	54,467									
2003	73	517	1,593	2,328										
2004	127	747	2,012											
2005	115	711												
2006	559													

A.1.2 Incurred claims

1993	2,609	6,730	9,659	12,084	12,801	13,458	13,569	13,552	13,414	13,461	13,471	13,481	13,584	13,569
1994	5,556	10,517	13,594	15,635	16,988	17,436	18,065	18,204	18,571	18,799	18,747	18,751	18,973	
1995	1,461	5,481	7,258	8,731	9,739	10,478	11,006	11,056	11,088	11,121	11,121	11,161		
1996	1,697	5,772	7,496	9,178	10,080	10,846	11,011	11,127	10,700	10,670	10,668			
1997	1,474	4,189	7,621	10,210	11,752	12,657	13,011	13,314	13,640	13,684				
1998	1,664	12,524	22,283	24,423	24,959	25,788	26,771	26,844	28,454					
1999	4,633	28,059	45,391	51,533	62,706	64,962	67,509	69,657						
2000	13,853	49,104	82,185	117,950	129,088	137,329	138,833							
2001	10,311	47,971	80,236	103,794	113,943	117,873								
2002	4,602	18,402	36,267	53,627	63,363									
2003	231	1,659	3,323	3,757										
2004	582	1,690	2,634											
2005	543	3,024												
2006	2,752													

A.1.3 Premium

1993	9,206	12,421	14,247	14,591	14,401	14,572	14,628	14,634	14,634	14,634	14,616	14,616	14,617	14,618
1994	4,966	9,290	11,293	11,593	11,491	11,812	11,967	11,967	11,967	11,968	11,968	11,968	11,913	
1995	4,023	8,167	9,164	8,358	8,259	8,281	8,285	8,286	8,286	8,286	8,288	8,419		
1996	3,513	8,283	8,208	8,015	8,029	8,028	8,044	8,041	8,041	8,046	8,045			
1997	2,956	6,897	8,586	8,677	8,609	8,646	8,656	8,674	8,661	8,668				
1998	3,935	8,984	13,778	13,757	13,366	13,381	13,378	13,381	13,281					
1999	14,593	29,079	32,777	33,364	33,268	33,307	33,266	33,316						
2000	16,428	47,209	52,678	53,624	53,673	53,541	53,499							
2001	28,913	63,532	65,714	66,977	66,763	67,086								
2002	22,951	46,264	48,791	48,835	48,814									
2003	5,563	8,918	9,394	9,143										
2004	3,889	6,630	8,301											
2005	4,893	7,919												
2006	11,643													

A.1.4 Premium rate index

		Basic Chain Ladder Ultimates		
Year	Q	Paid	Incurred	Premium
1993	1.170	13,551	13,569	14,618
1994	0.682	19,069	18,952	11,914
1995	0.816	11,305	11,261	8,402
1996	0.814	10,724	10,777	8,059
1997	0.684	13,317	13,813	8,680
1998	0.533	22,852	28,859	13,303
1999	0.507	67,323	71,977	33,313
2000	0.402	145,197	145,963	53,535
2001	0.581	126,142	126,812	67,199
2002	0.745	69,064	71,432	49,075
2003	2.136	3,682	4,737	9,153
2004	2.341	4,893	4,328	8,380
2005	1.184	3,257	8,225	8,847
2006	1.405	12,207	29,675	27,248

APPENDIX B – MEAN DELAYS BY DEVELOPMENT YEAR

This appendix derives the approximations given in Table 2 (Section 2.3.3) which is reproduced below for convenience. The values in the table are approximate mean delays (in years) from loss occurrence to end of development year. For accident year cohorts, development year 0 is the year in which the loss occurs. For underwriting year cohorts, development year 0 is the year of policy inception (that is, the year in which the cover provided by a policy commences).

Table 2: Approximate mean delay in each development year

Development year (*d*)	0	1	2	3	4	5+
Accident year mean delay (*t*)	0.5	1.5	2.5	3.5	4.5	d+0.5
Underwriting year mean delay (*t*)	0.333	1	2	3	4	d

The figures in Table 2 for accident year cohorts follow immediately from an assumption that losses occur uniformly throughout the accident year. The mean delay until the end of development year zero (which is the accident year itself) is then obviously half a year. The other values in the table are equally obvious for accident years.

For underwriting years it is assumed that:

(a) Policies incept uniformly throughout the year.

(b) Policies are in force for one year.

(c) Accidents occur uniformly throughout the year of cover provided by each policy.

For development year 1 we aim to find the mean delay between the accident date and the end of development year 1. Development year 1 is the year following the underwriting year. By assumptions (a) and (b) policies expire uniformly throughout development year 1, and all covered losses will have occurred by end of development year 1. By assumptions (b) and (c) the mean accident date on a policy is half a year after policy inception. By assumption (a) the mean point of policy inception is midway through the underwriting year. Therefore, over all policies, the mean accident date is one year after the start of the underwriting year, that is, at the end of development year zero. So the mean delay since accident occurrence at the end of development year 1 is $t = 1$. Clearly, for all later development years ($d > 1$) the mean delay to end of development year d is $t = d$.

For development year 0 the situation is more complex because only half of all covered losses are expected to have occurred by the end of development year 0 (because, by assumption (a), the mean policy inception date is 0.5 years before the end of development year zero). Because of this, the expected proportion of ultimate U that will be paid by end of development year zero is approximately 0.5 * $F_P(t)$ (instead of $F_P(t)$ for other development years) where t is the mean delay between accident date and end of development year 0 for the 50% accidents that occur before the end of development year 0.

Instead of explicitly including the factor 0.5 in the model for underwriting year cohorts, a factor is estimated by least squares as described in Section 2.5.3.

To find the mean delay t for the 50% of accidents expected to occur before the end of development year 0, we use s to denote the inception date of a policy: $s = 0$ corresponds to an inception date at the start of the underwriting year, and $s = 1$ corresponds to an inception date at the end of the underwriting year.

Since all policies are in force immediately before the end of the underwriting year, a delay $t = 0$ is possible on all policies regardless of the value of s. At the other extreme, a delay $t = 1$ is possible only on policies incepting at the start of the underwriting year (that is, on policies with $s = 0$). In general, for t between 0 and 1, a delay t is possible only on policies with $s < (1\text{-}t)$. So the mean delay is the weighted average of all values of t from 0 and 1, with weights proportional to $(1\text{-}t)$. That is:

$\text{Mean}(t) = \int t.(1-t).dt \,/\, \int (1-t).dt$ where both integrals are from $t = 0$ to $t = 1$.

Evaluating these integrals gives: $\text{Mean}(t) = (1/6)/(1/2) = 1/3$.

Abbreviations and notations

The table below gives an alphabetical list of all abbreviations and notation used in the paper. Items marked * in the second column are items of data. All other quantities are calculated from the data items. The final column shows subscripts that are sometimes applied to the symbol given in the first column:

- P/R means a subscript is used to distinguish parameters relating to payment and reporting delays.
- I/P/X means a subscript is used to distinguish quantities relating to incurred, paid, and exposure data.
- *j* means this subscript is sometimes applied to distinguish values relating to different origin years.

Symbol	**Data**	**Represents**	**Subscripts**
a		shape parameter of cumulative development curve	P/R and j
a_0, a_1		parameters linking a to Q_j (as in Equations 6 and 7)	P/R
b		mean case reserve bias factor	j
β_0, β_1		parameters linking b to Q_j (see Equation 6)	
β_2		parameter linking b to development time (see Equation 6a)	
BF		abbreviation for Bornheutter-Fergusson	
c		shape parameter of cumulative development curve	P/R and j
γ_0, γ_1		parameters linking c to Q_j (as in Equations 6 and 7)	P/R
CL		abbreviation for chain ladder	
d	*	development period in run-off array: $d = 0, 1, 2\ldots$	
$F_I(t)$		cumulative incurred run-off curve	j
$F_P(t)$		cumulative paid run-off curve	j
$F_R(t)$		cumulative distribution of reporting delays	j
θ_0, θ_1		adjustments to cumulative development in years 0 and 1 (Section 2.5.3)	I/P
$\Gamma(.)$		the Gamma function of mathematics	
I_{jd}	*	cumulative incurred development data	
j	*	origin year: $j = 1, 2, \ldots J$	
J	*	number of origin years in run-off array	
n_I	*	number of observations in incurred run-off array (I_{jd})	
n_P	*	number of observations in paid run-off array (P_{jd})	
n_X	*	number of origin years with known exposure (X_j or Prem$_j$)	
p_I		number of parameters estimated from incurred data (I_{jd})	
p_P		number of parameters estimated from paid data (P_{jd})	
p_X		number of parameters estimated from exposure data (X_j or Prem$_j$)	
P_{jd}	*	cumulative paid development data	
$Prem_j$	*	ultimate premium for origin year j	
Q_j	*	premium rate index for origin year j	
r		expected ultimate loss per unit of exposure	j
R		expected ultimate loss ratio (i.e., ultimate loss per unit of premium)	j
ϱ_0, ϱ_1, ϱ_2		parameters in model for r and R (see Equation 14)	
RSS		residual sum of squares (that is, sum of squared residuals)	$I/P/X$
s		scale parameter of cumulative development curve	P/R and j
σ_0, σ_1		parameters linking s to Q_j (see Equation 7)	P/R
σ_I^2		typical variance of a cumulative incurred observation (I_{jd})	
σ_P^2		typical variance of a cumulative paid observation (P_{jd})	
σ_X^2		ratio of variance to mean for an ultimate loss amount (U_j)	
t		continuous development time	
U_j		ultimate cumulative loss for origin year j	
ULR		abbreviation for ultimate loss ratio	
w_I		weight of incurred RSS relative to paid RSS in least squares estimation	
w_X		weight of exposure RSS relative to paid RSS in least squares estimation	
X_j	*	exposure for origin year j	

Biography of the Author

Thomas Wright is a consulting actuary who has worked mainly in non-life insurance since 1988, and with Deloitte & Touche LLP in London since 2001. He is a Fellow of the Institute of Actuaries, a Fellow of the Royal Statistical Society, and a Chartered Statistician. He is the author of several papers and articles on stochastic claims reserving and he was one of the pioneers of generalized linear modeling in personal lines rate-making. In 1991 he co-authored "Statistical Motor Rating: Making Effective Use of Your Data," which was published by the UK Institute of Actuaries and has since become a standard reference work. He was joint winner of the CAS/COTOR challenge in 2005 and is currently an active member of the Working Party on Reserve Uncertainty of the Institute of Actuaries General Insurance Reserving Oversight Committee (GIROC).

thwright@deloitte.co.uk

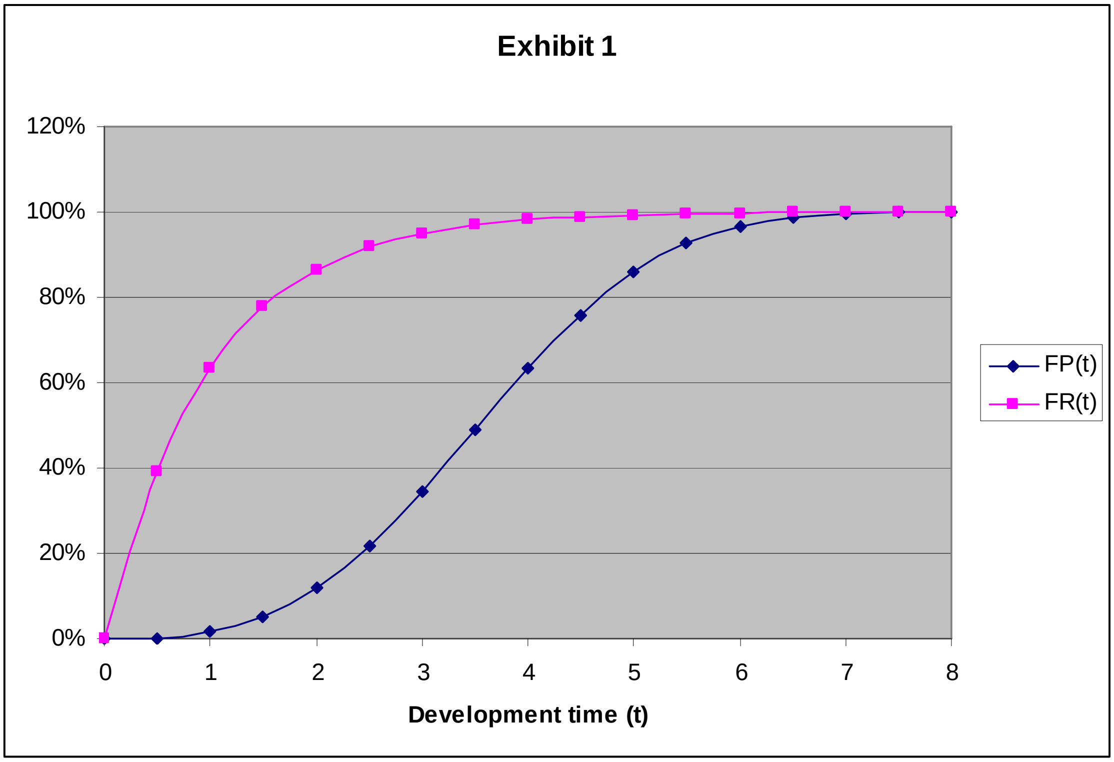
Exhibit 1
120%
100%
80%
60%
40%
20%
0%
0
1
2
3
4
5
6
7
8
Development time (t)
FP(t)
FR(t)

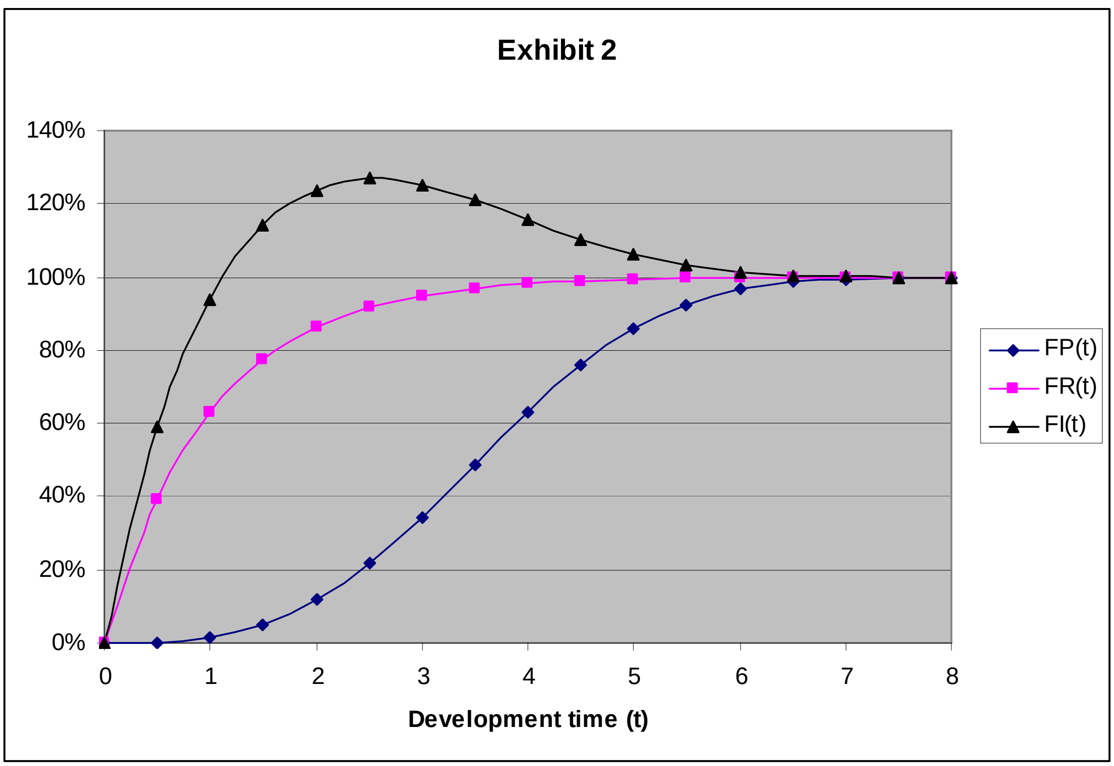
Exhibit 2
140%
120%
100%
80%
60%
40%
20%
0%
0
1
2
3
4
5
6
7
8
Development time (t)
FP(t)
FR(t)
FI(t)