BALANCING DEVELOPMENT AND TREND IN LOSS RESERVE ANALYSIS

SPENCER M. GLUCK

Abstract

The most common loss reserving procedures emphasize development-based projections, with implied trends examined for reasonableness and considered on an ad-hoc basis. This paper presents relatively simple methods for reflecting development and trend simultaneously, with weights that reasonably reflect the relative accuracy of the two types of projections. The Stanard–Bühlmann or "Cape Cod" method is a special case of these methods, which are denoted "Generalized Cape Cod" methods. The Appendices present underlying variance structures under which the weights used in the Generalized Cape Cod methods are optimal.

1. INTRODUCTION AND OVERVIEW

Commonly applied actuarial procedures involve projections in two "directions" of the traditional loss development triangle:

1. The development direction

We use the term "development" to refer to the emergence of information for a single year of origin. Development projections involve the measurement, selection, and application of development patterns. While the measurement and selection of development patterns often involve data from multiple years of origin, the application of the development pattern is made to each year of origin independently.

2. *The trend direction*

We use the term "trend" to refer to projections made using relationships among amounts for different years of origin. Trends, as used herein, refer to expected changes in the ratio of a projected amount to an exposure base. For simplicity, we will generally refer to the projected amount as "losses," the exposure base as "exposures" and the ratio of losses to exposures as "pure premiums."

In fact, the methods presented are more general, and have much wider potential application. For example, if the projected amount is losses, potential exposure bases include ultimate claim counts and premiums, in which case the quantity denoted "pure premium" herein would really be severity or loss ratio, respectively. Other examples of potential combinations of projected amounts and exposure bases are listed in Section 9.

In reserving methodology, primary emphasis is often given to development projections, with implied trends perhaps examined for reasonableness and ad hoc modifications sometimes made to development projections, particularly for recent years of origin.

The Bornhuetter–Ferguson method is commonly used to blend a development projection with an "a priori" result. While that a priori result may well be based on a trend projection of some kind, the basis for the a priori result is, in general, unspecified [1]. The Stanard–Bühlmann or "Cape Cod" method is an application of the Bornhuetter–Ferguson method in which the a priori result comes from a specified, trend-based calculation.[1] Regression-based or dynamic stochastic models can be used to reflect the development and trend directions simultaneously, but these models are not in widespread use.[2]

[1]See Stanard [2], Bühlmann [3] and Patrik [4].

[2]See, for example, DeJong and Zehnwirth [5], Taylor [6] and Wright [7].

The goal of this paper is to present methods that simultaneously reflect development and trend in a unified approach, reasonably reflecting the relative accuracy of the two types of projections. Additional goals are that the methods be practical, accessible to most practicing actuaries, and easily integrated with the most common actuarial procedures and the types of data that are readily available.

The methods discussed all estimate ultimate losses in a two step process; first, the expected losses for each year of origin are estimated based on a weighted average of the results from all years, developed and trended as appropriate; then the expected losses and actual losses for each year of origin are weighted together using the Bornhuetter–Ferguson (or similar) method. The Stanard–Bühlmann or Cape Cod method is a special case of this general approach, and we refer to the broader family of methods as Generalized Cape Cod Methods.

After introducing preliminary notation in Section 2, the paper proceeds as follows:

Section 3: The Bornhuetter–Ferguson method is presented along with a statistical justification for the Bornhuetter–Ferguson weights.

Section 4: The general framework for calculating expected losses as a weighted average of all years' results is presented, along with a discussion of variance relationships that should be reflected in the weights.

Section 5: The Stanard–Bühlmann or Cape Cod Method is presented and is shown to fit the general framework of Section 4. Two potentially significant shortcomings of the method are identified.

Sections 6 and 7: Two generalizations of the Cape Cod method are presented, designed to overcome the shortcomings identified in Section 5.

Section 8: The Bornhuetter–Ferguson calculation is revisited in conjunction with the expected loss estimates presented in Section 5 through 7.

Section 9: A number of potential applications of the methodology are listed.

Section 10: Conclusion

A glossary of notation is provided at the end of the paper including preliminary notation from Section 2, plus additional notation introduced in other Sections and Appendices.

The Appendices provide additional details of simple variance models consistent with the methods presented. Using these models, calculations are presented that use the data triangle to assist in the selection of parameters for Section 6 and Section 7 models.

The paper is organized with mathematics of any length or complexity consigned to the Appendices, and it is intended that the body of the paper can be read without the Appendices. Furthermore, while the Appendices provide calculations for estimating certain model parameters, the procedures of the paper will provide reasonable and useful results with judgmentally selected values for these parameters.

2. AVAILABLE DATA AND NOTATION

The following are presumed to be available (with $i = 1 \ldots N$):

	Notation
• The current evaluation of losses for each year of origin i	LTD_i
• Cumulative development factors appropriate to project losses to their ultimate value (note that the subscript refers to year of origin rather than maturity)	DF_i

- A measurement of the relative exposure per year of origin — E_i
- Trend factors to adjust for the change in expected losses per exposure from year of origin i to year of origin j — TF_{ij}

Additional Notation:

- Ultimate losses for year of origin i — ULT_i
- Thus, $LTD_i \times DF_i$ is an estimate of ULT_i
- $ULT_i \div E_i$ (i.e. pure premium) — PP_i
- Expected Value — E( )
- Variance — Var( )

The carat (^) is used to denote estimation; i.e., a quantity with a hat over it is an estimate of the quantity beneath the hat.

The derivation of the factors DF_i and TF_{ij} is not addressed in this paper. It is presumed that the actuary has applied appropriate calculations, adjustments, and judgments in selecting the factors so that $LTD_i \times DF_i$ is the best available development estimate of ULT_i, and TF_{ij} is the best available estimate of $\mathrm{E}(PP_j) \div \mathrm{E}(PP_i)$.[3]

For the most part, the above information is presumed to constitute all of the available information. In addition, calculations are presented in the Appendices that use the underlying data triangle to estimate certain model parameters.

Additional notation is introduced at later points in the paper. For convenience, a glossary containing all notation is included at the end of the paper.

3. THE BORNHUETTER–FERGUSON METHOD

The Bornhuetter–Ferguson method is the most commonly used approach to blending development and trend projections if

---

[3]Some of the types of adjustments that may be necessary are discussed in Berquist and Sherman [8].

trended values from other years of origin are the basis for the estimate of expected ultimate losses. In the Bornhuetter–Ferguson method, ultimate losses are estimated as follows:

$$\hat{ULT}_i = LTD_i + \left(1 - \frac{1}{DF_i}\right) \times \hat{\mathrm{E}}(ULT_i) \tag{3.1}$$

where the source of the estimate $\hat{\mathrm{E}}(ULT_i)$ is unspecified. Expanding the first term, we have:

$$\hat{ULT}_i = \left(\frac{1}{DF_i}\right) \times LTD_i \times DF_i + \left(1 - \frac{1}{DF_i}\right) \times \hat{\mathrm{E}}(ULT_i) \tag{3.2}$$

and the Bornhuetter–Ferguson estimate is seen to be a weighted average of the development based estimate of ULT_i and $\hat{\mathrm{E}}(ULT_i)$. The weights are optimal[4] under the following constraints:

1. Expected losses are known (i.e. $\hat{\mathrm{E}}(ULT_i) = \mathrm{E}(ULT_i)$);
2. Unemerged losses are independent from the emerged losses;
3. The DF_is are known; and
4. For a given year of origin i, the variance of the development-based estimate of ultimate losses (i.e. $LTD_i \times DF_i$) is proportional to the development factor DF_i.

Proof of the above statement is provided in Appendix A.

In practice, Constraint 1 is obviously not met; the majority of this paper concentrates on producing the best possible estimate of $\mathrm{E}(ULT_i)$ using all of the available information per Section 2. Section 8 and Appendix E deal with the implications of eliminating this constraint, and it is demonstrated that the same weights

[4]Optimal weights are defined as those that produce the best (i.e. the minimum variance) linear unbiased estimate, given that the individual estimates being weighted together are themselves considered to be unbiased.

remain optimal if the estimate $\hat{\mathrm{E}}(ULT_i)$ is determined using the techniques of this paper.

Constraint 2 is assumed to hold in both the Bornhuetter–Ferguson method and in underlying variance models developed in Appendix B. The independence assumption is modified for the model in Section 7 and Appendix C.

Constraint 3 is assumed to hold throughout.[5] Given the imperfection of this assumption, results described as optimal should be considered only approximately optimal.

Constraint 4 will subsequently be denoted as the "Cape Cod variance assumption." This assumption, along with several other assumptions, and the "Cape Cod variance model" is presented in Appendix B. Relaxation of this constraint and the use of an alternative variance model is addressed in Section 7 and Appendix C.

4. A FRAMEWORK FOR ESTIMATING EXPECTED LOSSES FOR A GIVEN YEAR

Using the available information and notation per Section 2, the expected pure premium for year i can be estimated based on the data from year j as follows:

$$_j\hat{\mathrm{E}}(PP_i) = \frac{LTD_j \times DF_j}{E_j} \times TF_{ji} \tag{4.1}$$

where the subscript on the left denotes that the estimate is based on data from year of origin j. Although we usually think of trend factors moving forward in time, note that j can also equal i or be greater than i.

Thus, the data from each year of origin j provide a different estimate $_j\hat{\mathrm{E}}(PP_i)$. If these estimates were independent, then

[5]Constraint 3 may be violated in practice. The DF_is are usually themselves random variables, which makes the mathematical properties of development estimates less than ideal. Stanard concluded that development estimates are not generally unbiased. See Appendix A of [2]. On the other hand, in Mack's model, development estimates can be unbiased [10].

the optimal estimate of $\hat{E}(PP_i)$ would be a weighted average of the estimates ${}_j\hat{E}(PP_i)$ with the weights inversely proportional to the variances of the estimation errors.[6,7] If the DF_is were known, such an independence assumption would be plausible. Given the methods normally used to estimate the DF_is, independence is unlikely. Nevertheless, we will attempt to develop weights roughly in inverse proportion to the relative variances of the estimation errors associated with the individual projections.

Differences among the variances of the estimation errors associated with the estimates ${}_j\hat{E}(PP_i)$ are generally related to the volume of the data and to the development and trending calculations as follows:

1. *Volume*

   All other things being equal, we normally expect that a larger volume of data produces a lower variance estimate of pure premium than a smaller volume of data. If we consider the loss data itself as the result of a random sample of size E_j, then the variance of the pure premium projection would be inversely proportional to E_j, and the indicated weight directly proportional to E_j. All models discussed herein assume that variance is inversely proportional to E_j and all weighting systems include E_j as an element of the weights.

2. *Development*

   All other things being equal, we normally expect that less mature data will produce higher variance estimates

---

[6] A common statistics result. See Rohatgi [9, p. 352].

[7] When the amount being estimated is an expected value, the variance of the estimation errors equals the variance of the estimate, and the two terms may be used interchangeably. When the amount being estimated is an actual value (i.e., the realization of a random variable), then the distinction between the variance of the estimation error and the variance of the estimate is important.

than more mature data. Thus, in a reasonable weighting system, the relative weight will increase with the maturity of the data.

3. *Trending*

   Given the imperfections in exposure measurement and trend estimation, the use of one year's data to estimate pure premium for another year would be expected to increase the variance of the estimation error as compared to using the data from the year itself. The relative variance would be expected to increase as the length of time between the years increases. This effect, which could be described as the deterioration in the value of information with time, is dealt with in many areas of actuarial practice.

While the general variance relationships discussed above will usually hold, they should not be considered absolute nor is the list necessarily exhaustive. There may be specific cases when one or more of the above relationships do not hold. Furthermore, the variances associated with the estimates ${}_j\hat{E}(PP_i)$ may come from sources that are not reflected in the above relationships, with some complex interactions among them. Limited to the practical goal of a reasonable and useful weighting system, this paper presents simple, practical models of the variance structure that reasonably account for the variance relationships listed above.

5. THE STANARD–BÜHLMANN OR CAPE COD METHOD

The Stanard–Bühlmann or Cape Cod method compared favorably with other loss reserving techniques in a study by Stanard [2]. Stanard also cites unpublished work by Bühlmann [3], who coined the name "Cape Cod." Patrik presents the method as a reinsurance reserving technique in the Foundations of Casualty Actuarial Science textbook, using the name Stanard–Bühlmann [4, pp. 352–354].

Stanard's presentation of the method assumed that exposures were equal for all years. The presentation below allows for varying levels of exposure, using the notation from Section 2. For clarity of presentation, we will omit the trend factor from the formulas in the remainder of the paper; it is assumed that losses and/or exposures have been adjusted for trend so that the pure premiums are expected to be equal for all years.

The expected pure premium is estimated as follows:

$$\hat{\mathrm{E}}(PP) = \frac{\sum_i LTD_i}{\sum_i [E_i/DF_i]}. \tag{5.1}$$

Note that $\hat{\mathrm{E}}(PP)$ is written without subscript, since the value is presumed to be equal for all years of origin. The expected pure premium thus calculated using the data from all available years is then used to calculate a priori expected losses in the Bornhuetter–Ferguson procedure.

Table 1 includes the trend adjustment and displays the calculation of the expected pure premium.

In Table 2, the expected pure premium is used in the Bornhuetter–Ferguson calculation.

It is instructive to expand Equation 5.1. Rewriting the numerator, we have:

$$\hat{\mathrm{E}}(PP) = \frac{\sum_i [(LTD_i \times DF_i/E_i) \times (E_i/DF_i)]}{\sum_i [E_i/DF_i]}. \tag{5.2}$$

In this form, the value of $\hat{\mathrm{E}}(PP)$ can be seen to be the weighted average of the developed projected pure premiums for each year ($LTD_i \times DF_i \div E_i$) with the weights equal to the values E_i/DF_i. Thus, the method falls into the general framework discussed in Section 4.

TABLE 1

COMPANY XYZ
WORKERS COMPENSATION STUDY
DATA AS OF 12/31/92

CAPE COD METHOD
CALCULATION OF EXPECTED PURE PREMIUM

(1) Accident Year	(2) E_i Exposures	(3) LTD_i Paid Losses @ 12/31/92 (000's)	(4) $TF_{i,1992}$ Trend Factor to 1992 @ 11%	(5) Trended Paid Losses @ 12/31/92 (3) × (4) (000's)	(6) DF_i Cumulative Paid Loss Development Factor	(7) E_i/DF_i (2)/(6)
1979	914	491	3.8833	1,907	1.1200	816
1980	1,203	385	3.4985	1,347	1.1312	1,063
1981	1,264	949	3.1518	2,991	1.1538	1,096
1982	1,372	769	2.8394	2,184	1.1769	1,166
1983	1,422	944	2.5580	2,415	1.2122	1,173
1984	1,502	909	2.3045	2,095	1.2624	1,190
1985	2,090	1,345	2.0762	2,792	1.3239	1,579
1986	2,338	1,298	1.8704	2,428	1.4175	1,649
1987	2,456	1,375	1.6851	2,317	1.5531	1,581
1988	2,617	2,086	1.5181	3,167	1.7053	1,535
1989	2,774	2,153	1.3676	2,945	1.9171	1,447
1990	3,021	2,265	1.2321	2,791	2.4865	1,215
1991	3,067	2,345	1.1100	2,603	3.4906	879
1992	3,428	1,186	1.0000	1,186	6.6569	515
Total	29,468	18,500		33,166		16,903

(8) Expected Pure Premium (at accident year 1992 level): 1.9621
(8) = (total Col. 5)/(total Col. 7)

The Cape Cod weights reflect two of the three variance relationships identified in Section 4. Volume is reflected by having the weights proportional to E_i. Development is reflected by having the weights inversely proportional to DF_i. Variance related to trending is not reflected.

TABLE 2

COMPANY XYZ
WORKERS COMPENSATION STUDY
DATA AS OF 12/31/92

CAPE COD METHOD
ESTIMATION OF ULTIMATE LOSSES
(BORNHUETTER–FERGUSON METHOD)

(1) Accident Year	(9) Expected Pure Premium (8)/(4)	(10) Expected Ultimate Losses/1,000 (2) × (9) (000's)	(11) Expected Unpaid Losses [1 − 1/(6)] ×(10) (000's)	(12) Estimated Ultimate Losses (3) + (11) (000's)
1979	0.5053	462	49	540
1980	0.5608	675	78	463
1981	0.6225	787	105	1,054
1982	0.6910	948	143	912
1983	0.7670	1,091	191	1,135
1984	0.8514	1,279	266	1,175
1985	0.9451	1,975	483	1,828
1986	1.0490	2,453	722	2,020
1987	1.1644	2,860	1,018	2,393
1988	1.2925	3,382	1,399	3,485
1989	1.4347	3,980	1,904	4,057
1990	1.5925	4,811	2,876	5,141
1991	1.7677	5,421	3,868	6,213
1992	1.9621	6,726	5,716	6,902
Total		36,849	18,819	37,319

Using weights inversely proportional to the DF_is makes the implicit assumption that the relative variance of the development-based pure premium estimates are proportional to the DF_is[8] (the Cape Cod variance assumption). This same assumption was previously listed as underlying the Bornhuetter–Ferguson method.

[8]This result is developed in Appendix B of [2].

This very simple variance model is often adequate to account for the decreasing reliability of projections as development factors increase—but not always. For example, incurred loss development factors will often approach unity well before all the losses are settled and all variance is eliminated. Incurred loss development factors less than unity provide an example where the Cape Cod variance assumption is clearly unreasonable.[9] Note that the Bornhuetter–Ferguson method also produces unreasonable results in this case.[10] This potential problem is addressed in Section 7.

The failure to reflect variance related to trending can be a serious shortcoming in practice. Practitioners have sometimes found that the Cape Cod method gives excessive weight to out of date results. The problem can be severe when a very long data base is used, as is often the case in reinsurance applications. The problem can be addressed to some degree by limiting the number of years entering the Cape Cod calculation. A less arbitrary approach is to specifically account for the relationship between variance and trending in the weighting scheme, as is presented in the following section.

6. ACCOUNTING FOR YEAR-TO-YEAR VARIANCE

We present two approaches for accounting for variance related to trending. The first uses an exponential decay factor, which is simple to apply and has proven practical in applications, although it is not based directly on a mathematical model. The decay factor approach is the one cited in most other sections of this paper. The second approach, using an additive "adaptive variance" term, is based on a specific mathematical model and is directly analogous to techniques used in dynamic stochastic modeling. With suitably

---

[9]The implication would be that the immature projected pure premium is more reliable (i.e., has lower variance) than the actual ultimate pure premium.

[10]Used with a development factor less than unity, the Bornhuetter–Ferguson method produces a projection outside of the range of the development result and the expected result.

chosen parameters, the two approaches produce similar results. The adaptive variance section can be skipped without substantial loss of continuity. The mathematical model underlying the adaptive variance approach is used in Appendix D to develop indicated values of the adaptive variance and of the approximately equivalent decay factor.

The Decay Factor Approach[11]

We account for the variance related to trending by introducing an exponential decay factor to the original Cape Cod weighting scheme. Equation 5.2 becomes:

$$\hat{\mathrm{E}}(PP_i) = \frac{\sum_j \left[\left(\frac{LTD_j \times DF_j}{E_j} \right) \times \left(\frac{E_j}{DF_j} \right) \times D^{|i-j|} \right]}{\sum_j \left[\left(\frac{E_j}{DF_j} \right) \times D^{|i-j|} \right]};$$

$$\text{where} \quad 0 \le D \le 1. \qquad (6.1)$$

The weights $(E_j/DF_j) \times D^{|i-j|}$ now reflect volume (via E_j), development using the Cape Cod variance assumption (via $1/DF_j$), and trending via the exponentially decaying weight $D^{|i-j|}$. The exponentially decaying weight has the required property that the relative weight decreases as the length of the trending period, $|i-j|$, increases. Note that the value $\hat{\mathrm{E}}(PP_i)$ now contains a subscript denoting the year of origin, since the weights will now shift for each year of origin, causing the values of $\hat{\mathrm{E}}(PP_i)$ to "drift."[12]

In Table 3, the example from Table 1 is re-worked with an annual exponential decay factor of 0.75. The calculation of

[11] Used for many years in various consulting reports [11].

[12] The drifting value of $\hat{\mathrm{E}}(PP_i)$ is roughly analogous to the techniques of dynamic stochastic modeling, where the various model parameters may be allowed to drift over time. See, for example DeJong and Zehnwirth [5] and Wright [7].

TABLE 3

COMPANY XYZ
WORKERS COMPENSATION STUDY
DATA AS OF 12/31/92

GENERALIZED CAPE COD METHOD WITH DECAY
CALCULATION OF EXPECTED PURE PREMIUM FOR ACCIDENT YEAR 1990
(USING A DECAY RATE OF 75%)

(1)	(2) E_i	(3) LTD_i	(4) $TF_{i,1992}$	(5) DF_i	(6)	(7) $D^{\lvert i-1990 \rvert}$	(8)
Accident Year	Exposures	Paid Losses @ 12/31/92 (000's)	Trend Factor to 1992 @ 11%	Cumulative Paid Loss Development Factor	Indicated Ultimate Trended Pure Premium (3) × (4) × (5)/(2)	Decay @ 75%	Total Weight Assigned to Indicated Ult. Pure Premium (2) × (7)/(5)
1979	914	491	3.8833	1.1200	2.3364	0.0422	34.467
1980	1,203	385	3.4985	1.1312	1.2665	0.0563	59.888
1981	1,264	949	3.1518	1.1538	2.7303	0.0751	82.256
1982	1,372	769	2.8394	1.1769	1.8730	0.1001	116.709
1983	1,422	944	2.5580	1.2122	2.0585	0.1335	156.586
1984	1,502	909	2.3045	1.2624	1.7607	0.1780	211.758
1985	2,090	1,345	2.0762	1.3239	1.7689	0.2373	374.626
1986	2,338	1,298	1.8704	1.4175	1.4719	0.3164	521.875
1987	2,456	1,375	1.6851	1.5531	1.4652	0.4219	667.133
1988	2,617	2,086	1.5181	1.7053	2.0635	0.5625	863.228
1989	2,774	2,153	1.3676	1.9171	2.0349	0.7500	1,085.233
1990	3,021	2,265	1.2321	2.4865	2.2970	1.0000	1,214.961
1991	3,067	2,345	1.1100	3.4906	2.9625	0.7500	658.984
1992	3,428	1,186	1.0000	6.6569	2.3031	0.5625	289.662
Total	29,468	18,500					6,337.366

(10) Expected Pure Premium for Accident Year 1990 (at 1992 AY level): 2.0675
(10) = Average of Col. 6 weighted by Col. 8

TABLE 4

COMPANY XYZ
WORKERS COMPENSATION COMPANY
DATA AS OF 12/31/92

GENERALIZED CAPE COD METHOD WITH DECAY
ESTIMATION OF ULTIMATE LOSSES
(BORNHUETTER–FERGUSON METHOD)

(1) Accident Year	(11) Expected Pure Premium @ AY 1992 Level	(12) Expected Pure Premium (11)/(4)	(13) Expected Ultimate Losses (2) × (12) (000's)	(14) Expected Unpaid Losses [1 − 1/(5)] × (13) (000's)	(15) Estimated Ultimate Losses (3) + (14) (000's)
1979	1.9586	0.5044	461	49	540
1980	1.9246	0.5501	662	77	462
1981	1.9676	0.6243	789	105	1,054
1982	1.9290	0.6794	932	140	909
1983	1.9019	0.7435	1,057	185	1,129
1984	1.8644	0.8090	1,215	253	1,162
1985	1.8397	0.8861	1,852	453	1,798
1986	1.8246	0.9755	2,281	672	1,970
1987	1.8511	1.0985	2,698	961	2,336
1988	1.9250	1.2680	3,318	1,372	3,458
1989	1.9915	1.4562	4,039	1,932	4,085
1990	2.0675	1.6781	5,069	3,031	5,296
1991	2.1399	1.9278	5,913	4,219	6,564
1992	2.1486	2.1486	7,365	6,259	7,445
Total			37,652	19,708	38,208

the expected pure premium for accident year 1990 is shown in Table 3.

The analogous calculation is then performed for other accident years, and the results are recorded in Column 11 of Table 4. Note the drifting values of the expected pure premium in Column 11. (The single value 1.9617 was used for all years in the Table 1 and Table 2 calculation).

This structure, using a decay factor between zero and unity, conveniently collapses to the original Cape Cod method when $D = 1$ and to the development method when $D = 0$. Thus, adding the decay factor produces a compromise between the Cape Cod method and the development method, with the degree of compromise controlled by the decay factor. The value of the decay factor should be a function of the variance associated with development projections compared with the variance associated with trend projections. In general, lower decay factors are appropriate for large data bases exhibiting stable development with higher decay factors for smaller data bases with more erratic development.

Given that using the decay factor produces a compromise between the Cape Cod and development methods, and that both the Cape Cod and development methods represent documented methodology, use of the decay factor will fall within the framework of documented methodology with any value of the decay factor between zero and unity, and it is reasonable that the decay factor may be judgmentally selected. Alternatively, the relative variances in the development and trend directions can be measured from the data triangle and used to aid in the selection of the decay factor.

Appendix B presents a variance model for data in a development triangle array, consistent with the Cape Cod variance assumption. Using that model, a method for using the data triangle to determine the indicated decay factor is developed in Appendix D.

In judgmentally selecting decay factors over many years in practice, we have generally used values ranging from 50% to 100%, with 75% as a "default" value. Estimates using the Appendix D methodology appear to confirm that this range is reasonable.

The Adaptive Variance Approach

This approach is justified by assuming that the unknown values $\mathrm{E}(PP_i)$ observe a simple random walk, i.e.,

$$\mathrm{E}(PP_i) = \mathrm{E}(PP_{i-1}) + d \qquad (6.2)$$

where d is a random "disturbance" with mean zero and variance ${}_d\sigma^2$. We refer to the value of ${}_d\sigma^2$ as the adaptive variance.

Denote the variance of the development-based estimate of PP_j as σ_j^2, i.e.

$$\mathrm{Var}(LTD_j \times DF_j \div E_j) = \sigma_j^2.$$

Then, it can be shown that the variance of the estimation error associated with using the development-based estimate of PP_j as an estimate of $\exp(PP_i)$ is as follows:

$$\mathrm{Var}(\mathrm{E}(PP_i) - LTD_j \times DF_j \div E_j) = \sigma_j^2 + {}_d\sigma^2 \times |i-j|. \qquad (6.3)$$

The Cape Cod weights in Equation 5.2 assume that σ_j^2 is directly proportional to DF_j and inversely proportional to E_j, i.e.

$$\sigma_j^2 = k \times DF_j / E_j \qquad (6.4)$$

for some proportionality constant k. Substituting in Equation 6.3, we have:

$$\mathrm{Var}(\mathrm{E}(PP_i) - LTD_j \times DF_j \div E_j) = \frac{k \times DF_j}{E_j} + {}_d\sigma^2 |i-j| \qquad (6.5)$$

and the indicated weights would be in inverse proportion to the variances in Equation 6.5. To calculate these weights requires estimates of both the adaptive variance ${}_d\sigma^2$ and the proportionality constant k.

The adaptive variance approach collapses to the original Cape Cod method when ${}_d\sigma^2 = 0$ and approaches the development method as ${}_d\sigma^2$ approaches infinity. Methods for estimating k and ${}_d\sigma^2$ are provided in Appendices B and D.

Selecting an Approach

Although the adaptive variance approach is more directly tied to a mathematical model, we generally prefer the decay factor approach since:

- the two approaches produce similar results;
- it is simpler to apply;
- it directly reflects the degree of compromise between the Cape Cod and development methods; and
- it is unitless, and thus is more amenable to judgmental selection, evaluation of reasonableness, and comparisons among different data bases.

7. GENERALIZING THE DEVELOPMENT VARIANCE ASSUMPTION

In each method presented thus far, the relative variances arising from development have been modeled using the Cape Cod variance assumption. While this simple assumption is often adequate, it is rather crude and is sometimes sufficiently inaccurate that the methods (including the Bornhuetter–Ferguson method in general) are unusable or of limited effectiveness.

Rather than specify a relationship between the development-related variance and the development factors, we address the issue more generally by introducing an additional "variance factor," VF_i, defined as follows:

$$VF_i = \frac{\mathrm{Var}[LTD_i \times DF_i]}{\mathrm{Var}(ULT_i)} \tag{7.1}$$

The Cape Cod variance assumption is the special case when $VF_i = DF_i$.

In each previously presented formula for estimating expected pure premiums, the VF_is replace the DF_is as an element of the weights. Thus, the original Cape Cod weighting scheme (Equation 5.2) becomes:

$$\hat{\mathrm{E}}(PP) = \frac{\sum_i \left[\left(\frac{LTD_i \times DF_i}{E_i} \right) \times \left(\frac{E_i}{VF_i} \right) \right]}{\sum_i \left[\left(\frac{E_i}{VF_i} \right) \right]}. \tag{7.2}$$

Using the decay factor, Equation 6.1 becomes:

$$\hat{\mathrm{E}}(PP_i) = \frac{\sum_j \left[\left(\frac{LTD_j \times DF_j}{E_j} \right) \times \left(\frac{E_j}{VF_j} \right) \times D^{|i-j|} \right]}{\sum_j \left[\left(\frac{E_j}{VF_j} \right) \times D^{|i-j|} \right]}. \tag{7.3}$$

After the expected pure premium is estimated, the final step of the reserving procedure has been the application of the Bornhuetter–Ferguson method; however, with the alternative assumption, an "alternative Bornhuetter–Ferguson" calculation is indicated. We modify Equation 3.2, replacing the weights based on DF_i with weights based on VF_i, as follows:

$$ULT_i = (1/VF_i) \times LTD_i \times DF_i + (1 - 1/VF_i) \times \hat{\mathrm{E}}(ULT_i). \tag{7.4}$$

In Table 5 the expected pure premium for accident year 1990 is calculated using a decay factor of 0.75 and variance factors in Column 8 different from the development factors in Column 5.

After performing the analogous calculation for other accident years, the remainder of the methodology is shown in Table 6.

Appendix C presents an alternative variance model consistent with using variance factors different from the development

TABLE 5

COMPANY XYZ
WORKERS COMPENSATION STUDY
DATA AS OF 12/31/92

GENERALIZED CAPE COD METHOD WITH DECAY AND ALTERNATIVE VARIANCE FACTORS CALCULATION OF EXPECTED PURE PREMIUM FOR ACCIDENT YEAR 1990 (USING A DECAY RATE OF 75%)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	E_i	LTD_i	$TF_{i,1992}$	DF_i		$D^{\lvert i-1990\rvert}$	VF_i	
Accident Year	Expo- sures	Incurred Losses @ 12/31/92 (000's)	Trend Factor to 1992 @ 11%	Cumu- lative Incurred Loss Devel- opment Factor	Indicated Ultimate Trended Pure Premium (3) × (4) × (5)/(2)	Decay Rate @ 75%	Variance Factors	Total Weight Assigned to Indicated Ult. Pure Premium (2) × (7)/(8)
1979	914	684	3.8833	1.0000	2.9061	0.0422	1.1200	34.467
1980	1,203	490	3.4985	1.0050	1.4321	0.0563	1.1312	59.888
1981	1,264	1,068	3.1518	1.0100	2.6897	0.0751	1.1538	82.256
1982	1,372	817	2.8394	1.0151	1.7164	0.1001	1.1769	116.709
1983	1,422	1,022	2.5580	1.0252	1.8848	0.1335	1.2122	156.586
1984	1,502	913	2.3045	1.0406	1.4577	0.1780	1.2624	211.758
1985	2,090	1,597	2.0762	1.0614	1.6838	0.2373	1.3239	374.626
1986	2,338	1,485	1.8704	1.0880	1.2926	0.3164	1.4175	521.875
1987	2,456	1,554	1.6851	1.1206	1.1948	0.4219	1.5531	667.133
1988	2,617	2,538	1.5181	1.1830	1.7417	0.5625	1.7053	863.228
1989	2,774	2,705	1.3676	1.2715	1.6957	0.7500	1.9171	1,085.233
1990	3,021	3,181	1.2321	1.4253	1.8491	1.0000	2.4865	1,214.961
1991	3,067	3,345	1.1100	1.7462	2.1140	0.7500	3.4906	658.984
1992	3,428	2,109	1.0000	2.2026	1.3551	0.5625	6.6569	289.662
Total	29,468	23,508						6,337.366

(11) Expected Pure Premium for Accident Year 1990 (at AY 1992 level): 1.6868

(11) = Average of Col. 6 weighted by Col. 9

TABLE 6

COMPANY XYZ
WORKERS COMPENSATION STUDY
DATA AS OF 12/31/92

GENERALIZED CAPE COD METHOD WITH DECAY AND ALTERNATIVE VARIANCE FACTORS ESTIMATION OF ULTIMATE LOSSES

(1) Accident Year	(11) Expected Pure Premium @ AY 1992 Level	(12) Expected Pure Premium (11)/(4)	(13) Expected Ultimate Losses (2) × (12) (000's)	(14) Development Basis Ultimate Losses (3) × (5) (000's)	(15) Estimated Ultimate Losses (13) × [1 − 1/(8)]+ (14)/(8) (000's)
1979	1.9586	0.5044	461	684	660
1980	1.9025	0.5438	654	492	511
1981	1.8916	0.6002	759	1,079	1,036
1982	1.8072	0.6365	873	829	836
1983	1.7450	0.6822	970	1,048	1,034
1984	1.6784	0.7283	1,094	950	980
1985	1.6377	0.7888	1,649	1,695	1,684
1986	1.5946	0.8525	1,993	1,616	1,727
1987	1.5873	0.9420	2,314	1,741	1,945
1988	1.6261	1.0712	2,803	3,002	2,920
1989	1.6557	1.2106	3,358	3,439	3,401
1990	1.6868	1.3690	4,136	4,534	4,296
1991	1.7071	1.5380	4,717	5,841	5,039
1992	1.6883	1.6883	5,787	4,645	5,616
Total			31,568	31,597	31,685

factors. Using that model, Equation 7.4 is demonstrated to be the indicated alternative to the Bornhuetter–Ferguson method.

It is beyond the scope of this paper to develop specific models of the relationship between variance and development. In practice, any reasonable variance factors will produce reasonable weights.

Measurement of the variance factors based on the actual data triangle is possible, but the available data may frequently be too limited to parameterize a model of any complexity. As a practical alternative, a "reference pattern" can be used, or a simple modification to a reference pattern can be made.

For the reference pattern to be useful, the values should be greater than unity as long as there is any significant remaining uncertainty in the development projection of ultimate losses. For example, if the Cape Cod variance assumption has been rejected for incurred development because the development factors decrease to unity (or less) faster than the uncertainty is eliminated, the paid development factors for the same business may provide a logical reference pattern (in the example of Table 5, the alternative variance factors are the paid development factors for the same business). A compromise between the paid and incurred development factors is another possible choice.

8. USING EXPECTED VALUE ESTIMATES IN THE BORNHUETTER–FERGUSON CALCULATIONS

The methodologies described in this paper estimate ultimate losses with a two step process: first, estimating expected ultimate losses by optimally combining information from all years; then using the estimated expected ultimate losses in the Bornhuetter–Ferguson or alternative Bornhuetter–Ferguson calculation. Proofs are provided in Appendices A and C that the Bornhuetter–Ferguson and alternative Bornhuetter–Ferguson weights are optimal, but the proofs are dependent on the constraint that the expected ultimate losses are assumed known.

In fact, the expected ultimate losses are not known. Rather, we are using an estimate of the expected ultimate losses. Furthermore, that estimate is not independent from the development result, since the development result from each year is part of the expected ultimate loss estimate.

In each of the estimates of expected ultimate losses presented in this paper, the expected ultimate loss estimate can be expressed as a weighted average of the development estimate from the year itself and other estimates independent of the data from the year itself (i.e., data from other years).

Thus:

$$\hat{\mathrm{E}}(ULT) = W' \times LTD \times DF + (1 - W') \times (Other)$$

where *Other* is an estimate of $\mathrm{E}(ULT)$, independent of *LTD* and *ULT*.

Additionally, the weights W' and $(1 - W')$ are inversely proportional to the variances of the estimates $LTD \times DF$ and *Other*, under the assumed variance models.

Appendix E addresses the issue of optimal Bornhuetter–Ferguson or alternative Bornhuetter–Ferguson weights, replacing the original assumption that $\mathrm{E}(ULT)$ is known with an assumption that the estimate $\hat{\mathrm{E}}(ULT)$ has the properties listed above. The result is that the exact same weights continue to be optimal.

9. APPLICATION

For convenience, we have referred to the quantity being projected by development methods as "losses," the exposure base as "exposures" and the ratio of the two as "pure premiums." However, there are many other potential applications. The methods described herein are useful any time we make a development-based projection and compare the result to some other predictive quantity. The following chart gives some examples:

QUANTITY BEING PROJECTED	"EXPOSURE" BASE	TREND ADJUSTMENT
Losses	Ratemaking Exposures	Pure Premium Trend
Losses	Ultimate Claim Counts	Severity Trend
Losses	Earned Premiums	Loss Ratio Index, or equivalently, Rate Adequacy Index
Claim Counts	Ratemaking Exposures	Frequency Trend
ALAE	Ultimate Losses	Expected Trend in ALAE/Loss Ratio (if any)
Salvage	Ultimate Losses	Expected Trend in Salvage/Loss (if any)
Excess Loss	Ultimate Limited Losses	Expected Trend in Excess/ Limited Losses (if any)

10. CONCLUSION

The techniques of this paper are useful in a wide variety of applications, and provide an alternative to the somewhat arbitrary judgments that are required when trend projections are incorporated only through reasonableness tests and ad hoc modifications to development projections.

The weighting methods presented herein are based on simplified variance structures designed to reasonably reflect the variance relationships that we typically expect to see. There is undoubtedly a good deal of room for improvement in this area, and the development of more rigorous variance models is an interesting and useful area for further research. However, the difference between reasonably good weights and optimal weights is often not significant, and the use of these techniques need not wait for improved variance models.

REFERENCES

[1] Bornhuetter, R. and R. Ferguson, "The Actuary and IBNR," *PCAS* LIX, 1972, pp. 181–195.

[2] Stanard, J., "A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques," *PCAS* LXXII, 1985, pp. 124–153.

[3] Bühlmann, H., "Estimation of IBNR Reserves by the Methods Chain Ladder, Cape Cod, and Complimentary Loss Ratio," unpublished, 1983.

[4] Patrik, G., "Chapter 6: Reinsurance," *Foundations of Casualty Actuarial Science*, Casualty Actuarial Society, 1990, pp. 277–374.

[5] DeJong, P. and B. Zehnwirth, "Claims Reserving, State-Space Models, and the Kalman Filter," *Journal of the Institute of Actuaries* 110, 1983, pp. 157–181.

[6] Taylor, G. C., "Regression Models in Claims Analysis I: Theory," *PCAS* LXXIV, 1987, pp. 354–383.

[7] Wright, T., "A Stochastic Method For Claims Reserving In General Insurance," *Journal of The Institute of Actuaries* 117, 1990, pp. 677–731.

[8] Berquist, J., and R. Sherman, "Loss Reserve Adequacy Testing: A Comprehensive, Systematic Approach," *PCAS* LXIV, 1977, pp. 123–184.

[9] Rohatgi, V. K., *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley & Sons, New York, 1976.

[10] Mack, T., "Measuring The Variability of Chain Ladder Reserve Estimates," *CAS Forum*, Spring 1994, pp. 101–182.

[11] Gluck, Spencer M., et al., various consulting reports, 1989–present.

GLOSSARY OF NOTATION

	Notation	Definition	Statistical Conception[13]
Section 2	N	Number of years of origin	
	LTD_i	Cumulative losses for year of origin i at current evaluation	Random variable
	DF_i	Cumulative development factor to ultimate applicable to losses for year of origin i at current evaluation	Treated as a known constant
	E_i	Measurement of relative exposure for year of origin i	Known constant
	TF_{ij}	Pure premium trend factor from year of origin i to year of origin j	Treated as a known constant
	ULT_i	Ultimate losses for year of origin i	Random variable
	PP_i	$ULT_i \div E_i$	Random variable
	E( )	Expectation	Operator
	Var( )	Variance	Operator
	$\hat{}$	Denotes estimation; i.e. the value is an estimate of the value under the "*hat*"	
Section 4	${}_j\hat{\mathrm{E}}(PP_i)$	Estimate of $\mathrm{E}(PP_i)$ based only on data from year of origin j. Defined in Equation 4.1	Estimated parameter, therefore, random variable
Section 5	PP	Single value of PP assumed to apply to all years of origin in the Cape Cod model	Random variable

[13]A number of random variables result from summarized data, and may be conceived of as sample sums or sample means (which are still random variables).

	Notation	Definition	Statistical Conception
Section 6	D	Usually used in the form $D^{\|i-j\|}$. First introduced in Equation 6.1. An exponential decay factor between zero and one, used to decrease relative weight as the length of the trend period increases	Unknown parameter, mostly treated as a known constant. If estimated, it is then an estimated parameter, therefore a random variable
	d	Random disturbance term, used to define a random walk in the values $E(PP_i)$. First introduced in Equation 6.2	Random variable
	${}_d\sigma^2$	Variance of d, called the "adaptive variance"	Unknown parameter
	k	A proportionality constant. First introduced in Equation 6.4. Also refer to Equation B.1.	Unknown parameter
Section 7	VF_i	"Variance factor" to reflect relative variances of development-based ultimate losses for different years of origin. Defined in Equation 7.1.	Treated as a known constant
Section 8	W'	Weight assigned to $LTD \times DF$ in an estimate of $E(ULT)$	
	Other	Estimate of $E(ULT)$ that is independent of the values LTD and ULT (normally from other years of origin)	
Appendix A	V^2	$Var(ULT)$	Unknown parameter
	W	Weight assigned to $LTD \times DF$ in Bornhuetter–Ferguson estimate of ULT	
	$\propto$	"Is proportional to"	
Appendix B	X_{ij}	Cumulative losses for year of origin i through development period j	Random variable

	Notation	Definition	Statistical Conception
	x_{ij}	Non-cumulative losses for year of origin i in development period j	Random variable
	P_j	Cumulative development pattern through period j	Treated as a known constant
	p_j	Non-cumulative development pattern in period j	Treated as a known constant
	n	Number of points of data	
	$\widehat{PP}_i$	Development based estimate of PP_i. Also serves as development based estimate of $E(PP_i)$	Estimated parameter, therefore, random variable
	$\widehat{PP}$	Cape Cod estimate of PP. Also serves as Cape Cod estimate of $E(PP)$. Defined in Equation B.5.	Estimated parameter therefore, random variable
	w_{ij}	Weight given to the value x_{ij}/E_ip_j in calculating $\widehat{PP}$. Defined in Equation B.6.	
	${}_D\widehat{PP}_i$	Estimate of $E(PP_i)$ using the Cape Cod with decay model, using decay factor D. Note a small inconsistency in the use of the "^" in that the value is an estimate of $E(PP_i)$ rather than PP_i	Estimated parameter therefore, random variable
	${}_Dw_{ij}$	The weight applied to the value x_{ij}/E_ip_j in calculating ${}_D\widehat{PP}_i$. Defined in Equation B.8.	
	$\hat{k}_D$	Estimate of k using decay factor D	Estimated parameter, therefore, random variable
Appendix C	LTD'	$LTD \times DF/VF$. Transformed value of LTD for use in the alternative variance model	Random variable

	Notation	Definition	Statistical Conception
	V_j	Cumulative "variance pattern" through development period j	Treated as known constant
	v_j	Non-cumulative variance pattern in development period $j (= V_j - V_{j-1})$	Treated as known constant
	X'_{ij}	$X_{ij} \times V_j / P_j$. Transformed value of X_{ij}	Random variable
	x'_{ij}	Non-cumulative transformed value $(= X'_{ij} - X'_{i,j-1})$	Random variable
Appendix D	ε_i	Random error related to estimate ${}_0\widehat{PP}_i$. Introduced in Equation D.1.	Random variable
	Δ_i	${}_0\widehat{PP}_i - {}_0\widehat{PP}_{i-1}$	Random variable
	Σ	Variance-covariance matrix of the vector Δ	Unknown parameter
	$\overline{\sigma}_i^2$	Variance of ${}_0\widehat{PP}_i$ for year of origin with average variance. Defined in Equation D.5.	

APPENDIX A

OPTIMALITY OF THE BORNHUETTER–FERGUSON CALCULATION

This Appendix provides a proof that the Bornhuetter–Ferguson weights are optimal (i.e., they produce the minimum variance estimate) under the constraints listed in Section 3. We will use the notation from Section 2, dropping the subscript denoting year of origin.

In addition, let $V^2 = \mathrm{Var}(ULT)$.

By Constraint 2, LTD is independent from $ULT - LTD$.

Therefore, $\mathrm{Cov}(ULT, LTD) = \mathrm{Var}(LTD)$.

By Constraint 4, $\mathrm{Var}(LTD \times DF) \propto DF$.

Noting that $DF = 1$ when $LTD = ULT$, the proportionality constant is $\mathrm{Var}(ULT)$, or V^2:

$$\mathrm{Var}(LTD \times DF) = V^2 \times DF$$

$$\mathrm{Var}(LTD) = V^2/DF$$

$$\begin{aligned}\mathrm{Var}(ULT - LTD) &= \mathrm{Var}(ULT) + \mathrm{Var}(ULT) - 2\,\mathrm{Cov}(ULT, LTD)\\ &= \mathrm{Var}(ULT) - \mathrm{Var}(LTD)\\ &= V^2(1 - 1/DF).\end{aligned}$$

The Bornhuetter–Ferguson method estimates ultimate losses as a weighted average of a development-based estimate of ultimate losses and expected ultimate losses, i.e.:

$$\widehat{ULT} = W \times LTD \times DF + (1 - W) \times \mathrm{E}(ULT); \qquad 0 \leq W \leq 1$$

Since the two estimates, $LTD \times DF$ and $\mathrm{E}(ULT)$, are clearly independent, optimal weights are inversely proportional to the variances of the estimation errors [9, p. 352]. Starting with the estimate $LTD \times DF$:

$$\begin{aligned}\mathrm{Var}(ULT - LTD \times DF) &= \mathrm{Var}(ULT) + \mathrm{Var}(LTD \times DF)\\ &\quad - 2\,\mathrm{Cov}(ULT, LTD \times DF)\\ &= V^2 + V^2 \times DF - 2 \times DF \times \mathrm{Var}(LTD)\\ &= V^2 + V^2 \times DF - 2V^2\\ &= V^2[DF - 1].\end{aligned}$$

The variance of the estimation error associated with using $\mathrm{E}(ULT)$ as an estimate of ULT:

$$\mathrm{Var}[ULT - \mathrm{E}(ULT)] = \mathrm{Var}(ULT) = V^2.$$

Calculating W in inverse proportion to the variances:

$$\begin{aligned}&= \frac{V^2}{V^2[DF-1] + V^2} = \frac{1}{DF - 1 + 1}\\ &= 1/DF\end{aligned}$$

which is the weight used in the Bornhuetter–Ferguson method.

APPENDIX B

THE CAPE COD VARIANCE MODEL

This Appendix provides a simple model for the variance structure of the development triangle consistent with the Cape Cod variance assumption.

We introduce additional notation to deal with the full data triangle. Also, note that whereas the DF_is in the prior notation carry a subscript denoting year of origin, the development patterns included in this Appendix carry subscripts denoting development period. Capital letters are used to denote cumulative values with lower case letters denoting the corresponding non-cumulative values.

All losses are presumed to have been trended to a common level.

Notation:

Cumulative Triangle Values: X_{ij}, $1 \leq i \leq N$, $1 \leq j \leq N$

Non-Cumulative Triangle Values: x_{ij} $(= X_{ij}$ for $j = 1$; $= X_{ij} - X_{i,j-1}$ for $j > 1)$

Cumulative Development Pattern: P_j, $(= 1/DF_{N-j+1})$

Non-Cumulative Development Pattern: p_j $(= P_j$ for $j = 1$; $= P_j - P_{j-1}$ for $j > 1)$

For year i, the number of points in X_{ij} is $N - i + 1$. Denote the total number of points in X_{ij} as $n = N(N + 1)/2$.

Additional Assumptions

- The values x_{ij} are assumed to be mutually independent. Within a given year of origin, this is somewhat more restrictive than (although clearly consistent with) the previous assumption of independence between the emerged and unemerged losses. Independence among values from different years of origin is an additional assumption briefly touched on in Section 4.

- The variance of PP_i is inversely proportional to E_i, i.e.

$$\mathrm{Var}(PP_i) = \frac{k}{E_i}. \tag{B.1}$$

This assumption was previously discussed in Section 4.

These additional assumptions are sufficient to determine that:

$$\mathrm{Var}(x_{ij}) = k \times E_i \times p_j. \tag{B.2}$$

Proof The Cape Cod variance assumption (using notation of this Appendix):

$$\mathrm{Var}\left(\frac{X_{ij}}{P_j}\right) \propto \frac{1}{P_j}.$$

Noting that $P_j = 1$ when $ULT_i = X_{ij}$, we have:

$$\begin{aligned}
\mathrm{Var}\left(\frac{X_{ij}}{P_j}\right) &= \frac{\mathrm{Var}(ULT_i)}{P_j}\\
\mathrm{Var}(X_{ij}) &= P_j \times \mathrm{Var}(ULT_i)\\
\mathrm{Var}(X_{i,j-1} + x_{ij}) &= \mathrm{Var}(X_{i,j-1}) + \mathrm{Var}(x_{ij})\\
\mathrm{Var}(x_{ij}) &= \mathrm{Var}(X_{ij}) - \mathrm{Var}(X_{i,j-1})\\
\mathrm{Var}(x_{ij}) &= P_j \times \mathrm{Var}(ULT_i) - P_{j-1} \times \mathrm{Var}(ULT_i)\\
\mathrm{Var}(x_{ij}) &= (P_j - P_{j-1}) \times \mathrm{Var}(ULT_i)\\
&= p_j \times \mathrm{Var}(ULT_i).
\end{aligned} \tag{B.3}$$

Noting that $ULT_i = E_i \times PP_i$ and using Equation B.1, we have:

$$\mathrm{Var}(ULT_i) = E_i^2 \times \mathrm{Var}(PP_i) = k \times E_i. \tag{B.4}$$

Substituting (B.4) in (B.3) produces Equation B.2.

The value k may be interpreted as the variance associated with one unit of exposure, when losses are fully developed. Each point, x_{ij} provides an independent estimate of $\mathrm{E}(PP_i)$, as $x_{ij}/E_i p_j$, with variance $k/E_i p_j$.

Optimality of the Development Estimates

Weighting all estimates of $\mathrm{E}(PP_i)$ from a given year of origin in inverse proportion to variances:

$$\hat{\mathrm{E}}(PP_i) = \frac{\sum_{j=1}^{N-i+1} (x_{ij}/E_i p_j) \times E_i p_j}{\sum_{j=1}^{N-i+1} E_i p_j}$$

$$= X_{i,N-i+1}/E_i P_{N-i+1}$$

which is the development estimate; call it $\widehat{PP}_i$.

Optimality of the Cape Code Estimate

We next assume that $\mathrm{E}(PP_i)$ is the same for all years i (we will write it as $\mathrm{E}(PP)$). Weighting all estimates of $\mathrm{E}(PP)$ in inverse proportion to variances:

$$\hat{\mathrm{E}}(PP) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N-i+1} (x_{ij}/E_i p_j) \times E_i p_j}{\sum_{i=1}^{N} \sum_{j=1}^{N-i+1} E_i p_j}$$

$$= \frac{\sum_{i=1}^{N} X_{i,N-i+1}}{\sum_{i=1}^{N} E_i \times P_{N-i+1}} \tag{B.5}$$

which is the estimate of the original Cape Cod method; call it $\widehat{PP}$.

Estimating the Proportionality Constant k

Estimates of the proportionality constant k are used in quantifying the decay factor (in Appendix D). The remainder of this Appendix deals with making these estimates. Noting that the estimate $\widehat{PP}$ is a weighted average of the individual estimates $(x_{ij}/E_i \times p_j)$, the value k can be estimated by averaging the sample variance estimates at each point:

$$\frac{\hat{k}}{E_i P_j} = \left(\widehat{PP} - \frac{x_{ij}}{E_i p_j}\right)^2$$

$$\hat{k} = \frac{(\widehat{PP} \times E_i p_j - x_{ij})^2}{E_i p_j}.$$

The sample variance estimates are biased low due to degrees of freedom of the estimate $\widehat{PP}$.[15] In a weighted average, the bias is different at each point. If a given point has weight w_{ij} such that $\sum w_{ij} = 1$, then the bias correction at that point is $1/(1 - w_{ij})$.

Define:

$$w_{ij} = \frac{E_i p_j}{\sum_{l=1}^{N} \sum_{j=1}^{N-l+1} E_l p_j} = \frac{E_i p_j}{\sum_{l=1}^{N} E_l p_{N-l+1}}. \tag{B.6}$$

Then the individual estimates of k, corrected for bias are:

$$\hat{k} = \frac{(\widehat{PP} \times E_i p_j - x_{ij})^2}{E_i p_j \times (1 - w_{ij})}.$$

[15] In keeping with our previously stated simplifying assumptions, we are treating the development pattern p_j and any trend factors used as known values. Since in practice these values are likely to be estimated from the x_{ij}, the measurement of variances is improved if the number of parameters in the development pattern is kept to the minimum necessary; thus, using a fitted curve for the development pattern is recommended.

Averaging all available estimates of k:

$$\hat{k} = \frac{\displaystyle\sum_{i=1}^{N}\sum_{j=1}^{N-i+1} \frac{(\widehat{PP} \times E_i p_j - x_{ij})^2}{E_i p_j (1 - w_{ij})}}{N}. \tag{B.7}$$

Estimating k Under the Cape Cod with Decay Model

The Cape Cod with Decay model allows that the value $E(PP_i)$ may vary among the years. Let ${}_D\widehat{PP}_i$ represent the estimate of $E(PP_i)$ using decay factor D. Using the notation of this section:

$$_D\widehat{PP}_i = \frac{\displaystyle\sum_{l=1}^{N}\sum_{j=1}^{N-l+1} (x_{lj}/E_l p_j) \times E_l p_j \times D^{|i-l|}}{\displaystyle\sum_{l=1}^{N}\sum_{j=1}^{N-l+1} E_l \times p_j \times D^{|i-l|}}$$

$$= \frac{\displaystyle\sum_{l=1}^{N} X_{l,N-l+1} \times D^{|i-l|}}{\displaystyle\sum_{l=1}^{N} E_l \times P_{N-l+1} \times D^{|i-l|}}$$

Note that ${}_1\widehat{PP}_i = \widehat{PP}$ for all i, and that ${}_0\widehat{PP}_i$ represents the development estimate for year i ($\widehat{PP}_i$).

Two modifications to Equation B.7 are indicated. First, the value $\widehat{PP}$ is replaced with the individual year values ${}_D\widehat{PP}_i$. Second, the degree of freedom correction is changed due to the change in weighting system. Denote the weight given to the point x_{ij} in calculating the value ${}_D\widehat{PP}_i$ as ${}_Dw_{ij}$. Then,

$$_Dw_{ij} = \frac{E_i p_j}{\displaystyle\sum_{l=1}^{N}\sum_{j=1}^{NJ-l+1} E_l \times p_j \times D^{|i-l|}} = \frac{E_i p_j}{\displaystyle\sum_{l=1}^{N} E_l \times P_{N-l+1} \times D^{|i-l|}}. \tag{B.8}$$

Note that if $D < 1$, the values ${}_Dw_{ij}$ are strictly greater than the values w_{ij}. Thus, the full set of values ${}_Dw_{ij}$ does not make a single set of weights. This is because when you change the subscript i, ${}_Dw_{ij}$ now refers to a weight in a different weighted average.

The formula for k reflecting decay factors less than one is as follows:

$$\hat{k}_D = \frac{\displaystyle\sum_{i=1}^{N}\sum_{j=1}^{N-i+1} \frac{({}_D\widehat{PP}_i \times E_i p_j - x_{ij})^2}{E_i p_j (1 - {}_Dw_{ij})}}{N}. \tag{B.9}$$

Note that in the special case when $D = 0$, $\hat{k}_0$ is based on observed within-year variance only. Also note that there will be no variance estimate available at the point $x_{N,1}$ (there will be no degrees of freedom). Thus, in the case of $D = 0$, Equation B.9 is modified by ending the first summation at $i = N - 1$ and changing the denominator to $N - 1$.

APPENDIX C

THE ALTERNATIVE VARIANCE MODEL

This Appendix provides an alternative variance model corresponding to the use of variance factors different from the development factors.

Justification for the Alternative Model

Given that our method is to weight together individual year development estimates that are assumed to be mutually independent, and that relative variances of those estimates are those assumed in Section 7, the weights in Section 7 follow directly.

However, in developing a consistent underlying model, we encounter the following difficulty: for the overall estimate to be optimal, each year's development estimate must be the optimal estimate based on the data from that year alone. In fact, it can be proven that the following three assumptions are irreconcilable:

1. optimality of the individual development estimates;
2. independence of the emerged and unemerged losses; and
3. variance factors different from the development factors.

We address the difficulty by changing the independence assumption. We will demonstrate that with the alternative assumption, the "alternative Bornhuetter–Ferguson" calculation (Equation 7.4) is indicated, that the development estimate is optimal for each individual year, and that the Cape Cod method with alternate variance factors is optimal using data from all years.

Note, however, that the alternative model was selected for mathematical convenience only. We will conclude this section with a brief discussion of whether the alternative model is intuitively reasonable.

Changing the Independence Assumption

We accomplish the alternative independence assumption by defining a transformation of the data. Independence assumptions are then assumed to hold for the transformed data, rather than the original data.

First, we use the cumulative notation of Appendix A.

Define $LTD' = LTD \times DF/VF$.

Thus $LTD' \times VF = LTD \times DF$.

We assume that LTD' is independent from $ULT - LTD'$.

Next, we present the alternative model in full triangle detail, using the notation of Appendix B.

In addition, we introduce the following notation:

Cumulative Variance Pattern: $V_j \; (= 1/VF_{N-j+1})$

Non-Cumulative Variance Pattern: $v_j \; (= V_j \text{ for } j = 1; \; = V_j - V_{j-1} \text{ for } j > 1)$

Let

$$X'_{ij} = X_{ij} \times V_j / P_j$$

$$x'_{ij} = X'_{ij} \qquad \text{for} \quad j = 1;$$

$$= X'_{ij} - X'_{i,j-1} \qquad \text{for} \quad j > 1$$

The values x'_{ij} are now assumed to be mutually independent.

Revisiting the Bornhuetter–Ferguson Calculation

We perform calculations analogous to those of Appendix A, with some changes to the underlying constraints. As in Appendix A, we drop the subscript denoting year of origin. Reviewing the underlying constraints:

1. Expected losses are known (i.e. $\hat{\mathrm{E}}(ULT) = \mathrm{E}(ULT)$).
2. Originally, unemerged losses $(ULT - LTD)$ were assumed to be independent from emerged losses (LTD). As defined previously, we now assume that $(ULT - LTD')$ is independent from (LTD').
3. Both development factors (DF) and variance factors (VF) are now assumed to be known.
4. The variance of $LTD \times DF$ is now assumed to be proportional to VF.

Let $V^2 = \mathrm{Var}(ULT)$. Then $\mathrm{Var}(LTD \times DF) = \mathrm{Var}(LTD' \times VF) = V^2 \times VF$

$$\mathrm{Var}(LTD') = V^2/VF$$

$$\mathrm{Var}(ULT - LTD') = V^2(1 - 1/VF).$$

Expressing the estimate of ULT as a weighted average of the development result and the expected ultimate losses:

$$\begin{aligned}\widehat{ULT} &= W \times LTD \times DF + (1 - W)\mathrm{E}(ULT)\\ &= W \times LTD' \times VF + (1 - W)\mathrm{E}(ULT).\end{aligned}$$

The remaining calculations, which are not shown here, exactly parallel those of Appendix A, except that LTD is replaced with LTD' and DF is replaced with VF. The indicated value of W is $1/VF$, which is the weight used in the alternative Bornhuetter–Ferguson calculation.

Optimality of the Development Estimates

We have previously defined mutually independent values x'_{ij}. An exactly analogous proof to that performed in Appendix B establishes that the variance of x'_{ij} is $k \times E_i \times v_j$.

Each value x'_{ij} now produces an independent estimate of $\mathrm{E}(PP_i)$, as $x'_{ij}/E_i v_j$ with variance $k/E_i v_j$.[16] Weighting all estimates for a given year of origin in inverse proportion to variances:

$$\hat{\mathrm{E}}(PP_i) = \frac{\sum_{j=1}^{N-i+1} (x'_{ij}/E_i v_j) \times E_i v_j}{\sum_{j=1}^{N-i+1} E_i v_j} = X'_{i,N-i+1}/E_i V_{N-i+1} = X_{i,N-i+1}/E_i P_{N-i+1}$$

or the development estimate.

Optimality of the Cape Cod Estimates with Alternate Variance Factors

If $\mathrm{E}(PP_i)$ is assumed to be equal for all years i, then the weighted average of all estimates of $\mathrm{E}(PP)$ from all years of origin is as follows:

$$\hat{\mathrm{E}}(PP) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N-i+1} (x'_{ij}/E_i v_j) \times E_i v_j}{\sum_{i=1}^{N} \sum_{j=1}^{N-i+1} E_i v_j}$$

$$= \frac{\sum_{i=1}^{N} X'_{i,N-i+1}}{\sum_{i=1}^{N} E_i V_{N-i+1}} = \frac{\sum_{i=1}^{N} (X_{i,N-i+1}/P_{N-i+1}) \times V_{N-i+1}}{\sum_{i=1}^{N} E_i V_{N-i+1}}$$

[16] A proof that x'_{ij}/v_j is an estimate of ULT_i is provided later in this Appendix.

which is the Cape Cod estimate with the alternate variance factors.

Estimating the Proportionality Constant k

The estimate of the proportionality constant k exactly parallels the calculations presented in Appendix B, except that X'_{ij} and x'_{ij} replace X_{ij} and x_{ij}, and V_j and v_j replace P_j and p_j.

Thus Equations B.6 through B.9 become:

$$w_{ij} = \frac{E_i v_j}{\sum_{l=1}^{N} E_l V_{N-l+1}} \tag{C.1}$$

$$\hat{k} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N-i+1} \frac{(\widehat{PP} \times E_i v_j - x'_{ij})^2}{E_i v_j \times (1 - w_{ij})}}{n} \tag{C.2}$$

$${}_Dw_{ij} = \frac{E_i v_j}{\sum_{l=1}^{N} E_l V_{N-l+1} D^{|i-1|}} \tag{C.3}$$

$$\hat{k}_D = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N-i+1} \frac{({}_D\widehat{PP}_i \times E_i v_j - x'_{ij})^2}{E_i v_j (1 - {}_Dw_{ij})}}{N}. \tag{C.4}$$

Note the discussion in Appendix B regarding modifying Equation C.4 for the special case when $D = 0$.

We now provide a proof that x'_{ij}/v_j is an unbiased estimate of ULT_i, assuming that each development result is an unbiased estimate of ULT_i, and that the values V_j are known (i.e. not random variables).

Assume, for all X_{ij}:

$$ULT_i = \mathrm{E}(X_{ij}/P_j) = \mathrm{E}(X'_{ij}/V_j).$$

Then,

$$\mathrm{E}\left(\frac{x'_{ij}}{v_j}\right) = \frac{\mathrm{E}(x'_{ij})}{v_j} = \frac{\mathrm{E}(X'_{ij}) - \mathrm{E}(X'_{i,j-1})}{v_j}$$
$$= \frac{V_j(ULT_i) - V_{j-1}(ULT_i)}{v_j}$$
$$= \frac{(V_j - V_{j-1})(ULT_i)}{v_j} = ULT_i.$$

Plausibility of the Alternative Model

Having replaced the assumption of independence between the emerged and unemerged losses, the alternative model implies dependence. It can be proven that when the variance factor is larger than the development factor, the alternative model implies negative correlation between emerged and unemerged losses. In this section, we discuss the plausibility of that result under two scenarios.

The first scenario is an incurred development projection, which is the most common situation in which variance factors different from the development factors will be needed. In this situation, the inclusion of the case reserves in the data may lead to variance factors higher than the development factors and a presumed negative correlation between emerged and unemerged losses. In this case, the sign of the dependence is logical: relative over-reserving of cases for a particular year of origin will lead to a high error on the emerged losses and a low error on the unemerged losses, and vice versa. Of course, we have not addressed whether the amount of dependence predicted by the alternative model is reasonable.

A paid loss development scenario provides a counter-example. Assume that there are no partial payments, that average claim size tends to grow with the lag to settlement, and that the coefficient of variation of the claim size distribution is constant.

These assumptions imply that the independence model is appropriate, and yet the variance factors will be different from the development factors. In this case the alternative model appears inappropriate. If the variance factors are correct, the generalized Cape Cod weights still produce the optimal combination of the individual year development projections; however, the individual year development projections do not represent the optimal combination of data from a particular year of origin.

APPENDIX D

ADAPTIVE VARIANCES AND DECAY FACTORS

This Appendix provides a method for calculating the indicated decay factor D. The approach first calculates the indicated adaptive variance under the random walk model, and then calculates the approximately equivalent decay factor.

Calculating the Adaptive Variance

This approach for calculating the adaptive variance was used by Wright [7].

Recalling that the development-based pure premium estimate for year i is denoted ${}_0\widehat{PP}_i$,

$$ {}_0\widehat{PP}_i = \mathrm{E}(PP_i) + \varepsilon_i. \tag{D.1}$$

The random walk model connecting the values $\mathrm{E}(PP_i)$ is:

$$\mathrm{E}(PP_i) = \mathrm{E}(PP_{i-1}) + d \qquad \text{for} \quad i = 2,3,\ldots \tag{D.2}$$

The error term ε_i and the random disturbance term d are presumed to have variances σ_i^2 and ${}_d\sigma^2$, respectively. For the purposes of this calculation, we will also assume that ε_i and d are normally distributed.

The adaptive variance ${}_d\sigma^2$ is the variance of the differences between the true (unknown) parameters $\mathrm{E}(PP_i)$, not the estimates ${}_0\widehat{PP}_i$. We measure the adaptive variance by observing the differences in the estimates ${}_0\widehat{PP}_i$, and correcting for the estimation errors ε_i.

Let

$$\Delta_i = {}_0\widehat{PP}_i - {}_0\widehat{PP}_{i-1} \qquad \text{for} \quad i = 2,3,\ldots$$

Then,

$$\Delta_i = (\mathrm{E}(PP_i) + \varepsilon_i) - (\mathrm{E}(PP_{i-1}) + \varepsilon_{i-1}) = d + \varepsilon_i - \varepsilon_{i-1}.$$

The variance-covariance matrix, Σ, of the vector Δ is given by:

$$\Sigma = \begin{bmatrix} {}_d\sigma^2 + \sigma_1^2 + \sigma_2^2 & -\sigma_2^2 & 0 & 0 \\ -\sigma_2^2 & {}_d\sigma^2 + \sigma_2^2 + \sigma_3^2 & -\sigma_3^2 & 0 \\ 0 & -\sigma_3^2 & {}_d\sigma^2 + \sigma_3^2 + \sigma_4^2 & -\sigma_4^2 \\ 0 & 0 & -\sigma_4^2 & \text{etc.} \end{bmatrix}$$

Δ is normally distributed, and the value $\Delta^T \times \Sigma^{-1} \times \Delta$ is chi-squared distributed with $N - 1$ degrees of freedom ($N - 1$ is the length of the vector Δ).

The expected value of the chi-squared distribution is $N - 1$, so an estimate of σ_d^2 is given by solving the equation:

$$\Delta^T \times \Sigma^{-1} \times \Delta = N - 1 \qquad \text{for} \quad {}_d\sigma^2, \tag{D.3}$$

which can be solved numerically, given that estimates of σ_i^2 are available. It is possible that the variances σ_i^2 may be large enough compared to the differences Δ_i that $\Delta^T \times \Sigma^{-1} \times \Delta < N - 1$ for all ${}_d\sigma^2 > 0$. In this case, there is no demonstrable random walk and we set ${}_d\sigma^2$ to zero (and the decay factor D to unity).

Under the alternative variance model of Appendix C (which includes the Cape Cod variance model as a special case),

$$\sigma_i^2 = \frac{k}{E_i V_{N-i+1}}. \tag{D.4}$$

Given an estimate of k, Equation D.4 can be applied and then Equation D.3 solved for ${}_d\sigma^2$.

To convert an estimate of ${}_d\sigma^2$ to the approximately equivalent decay factor, we have used the following formula:[17]

$$D = \frac{\bar{\sigma}_i^2}{\bar{\sigma}_i^2 + {}_d\sigma^2}, \qquad \text{where} \quad \bar{\sigma}_i^2 = \frac{Nk}{\sum_{i=1}^{N} E_i V_{N-i+1}}. \tag{D.5}$$

[17]The above formula equates the decay factor and adaptive variance approaches at a lag of one year, for a year of origin with average variance.

To estimate k, we use the formulas of Appendix B or C; however, the derivation of $\hat{k}$ is dependent on an estimate of D. For the first iteration, we assume $D = 0$ and use Equation B.9 or C.4 to calculate $\hat{k}_0$.[18] We then apply Equations D.4, D.3, and D.5 above to estimate D. Equation B.9 or C.4 can then be used to estimate $\hat{k}_D$, and D can be re-estimated. This process can be applied repeatedly.

[18] $D = 0$ is a logical starting point since $\hat{k}_0$ is based entirely on within-year variance and provides an unbiased estimate of k regardless of the appropriate value of D. Given that $D > 0$, $\hat{k}_D$ is a superior (i.e. lower variance) estimate of k.

APPENDIX E

OPTIMALITY OF BORNHÜETTER–FERGUSON CALCULATION WITH RELAXED CONSTRAINTS

Appendices A and C provide proofs that the Bornhuetter–Ferguson weights are optimal based on the Cape Cod variance model and alternative variance model, respectively, with the additional constraint that the expected ultimate losses are known (i.e., $\hat{E}(ULT) = E(ULT)$).

The Bornhuetter–Ferguson and alternative Bornhuetter–Ferguson weights remain optimal using an estimate, $\hat{E}(ULT)$, if:

$$\hat{E}(ULT) = W' \times LTD \times DF + (1 - W')(Other) \tag{E.1}$$

where *Other* is an estimate of $E(ULT)$, independent of *LTD* and *ULT*, and

$$\frac{\mathrm{Var}(Other)}{\mathrm{Var}(LTD \times DF)} = \frac{W'}{(1 - W')} \tag{E.2}$$

i.e., $LTD \times DF$ and *Other* are weighted in inverse proportion to the variances of the estimates. All of the estimates of $E(ULT)$ described in this paper meet these conditions under the assumed variance models.

We provide the proof using the alternative variance model of Appendix C, which includes as a special case the original Cape Cod variance model.

Let $V^2 = \mathrm{Var}(ULT)$.

$$\mathrm{Var}(LTD \times DF) = V^2 \times VF \tag{E.3}$$

$$\widehat{ULT} = W \times LTD \times DF + (1 - W)\hat{E}(ULT) \tag{E.4}$$

Substituting Equation E.1 in Equation E.4:

$$\widehat{ULT} = W \times LTD \times DF + (1 - W)(W' \times LTD \times DF + (1 - W')(Other))$$

$$= (W + (1 - W)W') \times LTD \times DF + (1 - W)(1 - W')(Other)$$

$$= W^* \times LTD \times DF + (1 - W^*)(Other),$$

where $W^* = W + (1 - W)W'$. (E.5)

This is a weighted average of two independent estimates of *ULT*. The variance of the estimation error associated with the first estimate,

$$\mathrm{Var}(ULT - LTD \times DF) = V^2(VF - 1),$$

is a result developed in Appendices A and C.

For the second estimate,

$$\begin{aligned}
&\mathrm{Var}(ULT - Other)\\
&\quad = \mathrm{Var}(ULT) + \mathrm{Var}(Other)\\
&\quad = V^2 + \mathrm{Var}(LTD \times DF)(W'/1 - W'), \quad \text{using E.2}\\
&\quad = V^2 + V^2 \times VF(W'/1 - W'), \quad \text{using E.3}\\
&\quad = V^2[1 + VF(W'/1 - W')].
\end{aligned}$$

Calculating $1 - W^*$ in inverse proportion to variances:

$$1 - W^* = \frac{V^2(VF - 1)}{V^2(VF - 1) + V^2(1 + VF(W'/1 - W'))}$$

$$= \frac{VF - 1}{VF - 1 + 1 + VF(W'/1 - W')} = \frac{VF - 1}{VF(1 + W'(1 - W')}$$

$$= \frac{VF - 1}{VF(1/1 - W')} = \frac{(VF - 1)(1 - W')}{VF}. \tag{E.6}$$

Substituting $1 - W^* = (1 - W)(1 - W')$ (see E.5)

$$(1 - W)(1 - W') = \frac{(VF - 1)(1 - W')}{VF}$$

$$= 1 - \frac{VF - 1}{VF} = \frac{1}{VF}$$

which is the weight used in the alternative Bornhuetter–Ferguson calculation.