"A Bayesian Generalized Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving," R. J. Verrall, July 2004

KATRIEN ANTONIO,* JAN BEIRLANT,† and Tom Hoedemakers‡

AIM OF THIS COMMENT

Professor Verrall nicely illustrates how Bayesian models can be applied to claims reserving within the framework of Generalized Linear Models (GLMs) and how they lead to posterior predictive distributions of quantities of interest. In this discussion we apply a Bayesian model in the context of discounted loss reserves. The outcomes of this approach are compared with results on approximations for the distribution of the discounted loss reserve when the run-off triangle is modelled by a GLM. These approximations are based on the concepts of comonotonicity and convex order and are explained in full details in Hoedemakers et al. (2003) (for lognormal claims reserving models) and Hoedemakers et al. (2004) (for claims reserving using GLMs). The comonotonicity approach has been shown to provide elegant solutions to various other actuarial and financial problems involving the distribution of a sum of dependent random variables (check www.kuleuven.ac.be/insurance-research papers for more illustrations). We realize that the Bayesian posterior predictive distribution is a very general workhorse, which takes into account all sources of uncertainty in the model formulation and is applicable to different statistical domains, whereas our approximations originate from a specific actuarial context. We want to illustrate however that the predictive distribution based on the comonotonic bounds provides results that are close to the results obtained via Markov Chain Monte Carlo (MCMC). The main advantage of the bounds is that several risk measures (e.g., VaRs and stop-loss premia), and TailVaRs can be calculated easily from it. Moreover, the use of the upper bound (see below) leads to a conservative estimate.

GENERALIZED LINEAR MODELS FOR CLAIMS RESERVING: THE LIKELIHOOD-BASED APPROACH

We use the notation from Verrall (2004). An insurer is interested in the aggregated value $\Sigma_{i=2}^{n} \Sigma_{j=n+2-i}^{n} C_{ij}$ corresponding with the future part of a classical run-off triangle (as shown in Table 1 of Verrall's paper). In a (likelihood-based) GLM framework this reserve will be predicted by

$$\text{reserve} = \sum_{i=2}^{n} \sum_{j=n+2-i}^{n} \hat{\mu}_{ij},$$

with $\hat{\mu}_{ij} = g^{-1}\,[(\mathbf{R}\hat{\vec{\beta}})_{ij}]$. Here $\mathbf{R}$ denotes the design matrix, $\vec{\beta}$ is the vector of regression parameters, and g the link function of the GLM that is used to model both observed data and future observations. $\hat{\vec{\beta}}$ denotes the maximum likelihood estimator of $\vec{\beta}$.

The reserve is a provision for future payments. Therefore (in this discussion) our estimated loss reserve reflects the time value of money and we consider the *discounted* loss reserve. Assume that the reserve will be invested such that an amount of 1 at time $j - 1$ becomes e^{Y_j} at time j. The discount factor for a payment of 1 at time i is then given by $e^{-(Y_1+Y_2+\cdots+Y_i)} := e^{-Y(i)}$. Let us assume for instance that

$$Y(i) = \left(\mu + \frac{\delta^2}{2}\right)i + \delta B(i),$$

where $B(i)$ is the standard Brownian motion. We extended the WinBugs programs from Appendix B of Verrall (2004) to incorporate this discounting process. The discounted reserve now becomes

* Katrien Antonio is a PhD student at the University Center for Statistics, de Croylaan 54, 3001 Heverlee, Belgium, Katrien.Antonio@econ.kuleuven.ac.be.

† Jan Beirlant is a Professor at the University Center for Statistics, de Croylaan 54, 3001 Heverlee, Belgium, Jan.Beirlant@wis.kuleuven.ac.be.

‡ Tom Hoedemakers is a PhD student at the Department of Applied Economics, Catholic University Leuven, Naamsestraat 69, 3000 Leuven, Belgium, and at the University Center for Statistics, Tom.Hoedemakers@econ.kuleuven.ac.be.

$$R := \sum_{i=2}^{n} \sum_{j=n+2-i}^{n} \hat{\mu}_{ij} e^{-Y(i+j-n-1)}$$
$$= \sum_{i=2}^{n} \sum_{j=n+2-i}^{n} \hat{\mu}_{ij}$$
$$\exp\,[-(\mu + \delta^2/2)(i + j - n - 1) - \delta B(i + j - n - 1)]. \qquad (1)$$

In any realistic model for the return process, R will be a sum of strongly dependent random variables. Because one can not rely on traditional risk theory, it becomes very hard or even impossible to compute the cumulative distribution function (cdf) of R analytically. For actuaries, however, the complete predictive distribution of the reserve is important because different risk measures can be calculated from it. As illustrated by Verrall (2004) (for GLMs) and in earlier work by (for instance) de Alba (2002) (for lognormal models) Bayesian techniques are useful in this area as they provide the posterior predictive distribution of the reserve. This discussion compares the results from Bayesian GLMs (as used in Verrall) with the results based on upper and lower bounds for R for which the cdf can be easily calculated.

COMONOTONICITY AND CONVEX ORDER APPLIED TO LOSS RESERVING

Let $S = X_1 + \cdots + X_n$ be an arbitrary sum of random variables. In many actuarial and financial problems (e.g., claims reserving) the distribution of such a sum of random variables is of interest (check Dhaene et al. [2002b] for various illustrations). In general, however, this cdf can not be obtained analytically because of the dependency structure involved in the stochastic vector $(X_1, \ldots, X_n)$. Kaas et al. (2000) and Dhaene et al. (2002a) propose a possible way out by considering upper and lower bounds for (the cdf of) S that allow explicit calculations of various actuarial quantities.

The original dependency structure can be replaced by a simpler one, which leads to a more ‘‘dangerous’’ random variable, but allows analytical calculations. Decisions (e.g., ‘‘Which amount of money should be held aside?’’) based on this random variable lead to a safe, conservative strategy. However, in the context of claims reserving, provisions will be held that are too high. Therefore it makes sense that one also considers a less ‘‘dangerous’’ random variable allowing analytical calculations of various risk measures.

Kaas et al. (2000) and Dhaene et al. (2002a) provide a mathematical foundation for these intuitive ideas. The concept ‘‘more/less dangerous’’ is defined by the notion of convex order. $X \leq Y$ in convex order if and only if $E[X] = E[Y]$ and $E[(X - d)_+] \leq E[(Y - d)_+]$ for all real d. Upper and lower bounds for S are derived for which the terms in the involved sums are comonotonic, meaning that they are all monotonic (increasing or decreasing) functions of the same random variable. Precisely the quality of comonotonicity makes it possible to compute various actuarial quantities in an analytical way.

Hoedemakers et al. (2003) extend the results mentioned above from ordinary sums of variables to sums of scalar products of independent random variables. This extension allows to derive a lower and upper bound for the discounted loss reserve R. Based on the concept of comonotonicity, they explain how the cdf of these bounds can be derived explicitly.

NUMERICAL ILLUSTRATION

Consider the run-off triangle in Table 1, taken from Taylor and Ashe (1983) and used in various other publications on claims reserving.

These data are modelled using a gamma GLM with logarithmic link function and the following structure for the linear predictor

$$\begin{aligned}\log(\mu_{ij}) = {} & \alpha_1 I(i = 1) + \alpha_2 I(i = 2, 3) \\ & + \alpha_3 I(i = 4) + \alpha_4 I(i > 4) \\ & + \beta_1 I(j > 1) + \beta_2 I(j > 4) \\ & + (j - 5)\beta_3 I(5 < j < 9) \\ & + 3\beta_3 I(j > 8)\end{aligned}$$

whereby $I(.)$ denotes an indicator function.

The discounting process from (1) (with $\mu = 0.08$ and $\delta = 0.11$) is incorporated in the WinBugs code for the gamma GLM. To enable comparisons with the results based on the comonotonic bounds, flat priors were used both for the row and column parameters of the linear predictor and for the scale parameter in the gamma

Table 1
Run-Off Triangle

	1	2	3	4	5	6	7	8	9	10
1	357,848	766,940	610,542	482,940	527,326	574,398	146,342	139,950	227,229	67,948
2	352,118	884,021	933,894	1,183,289	445,745	320,996	527,804	266,172	425,046	
3	290,507	1,001,799	926,219	1,016,654	750,816	146,923	495,992	280,405		
4	310,608	1,108,250	776,189	1,562,400	272,482	352,053	206,286			
5	443,160	693,190	991,983	769,488	504,851	470,639				
6	396,132	937,085	847,498	805,037	705,960					
7	440,832	847,631	1,131,398	1,063,269						
8	359,480	1,061,648	1,443,370							
9	376,686	986,608								
10	344,014									

model. Table 2 contains the results for the 95th percentile of the predictive distribution obtained via MCMC simulations with the WinBugs program. A burn-in of 10,000 iterations was allowed, after which another 10,000 iterations were performed.

As mentioned before, an upper (denoted by R_u) and lower bound (denoted by R_l) for R (in convex order) can be derived. R_u and R_l are sums of comonotonic terms. Precisely this quality enables the analytical calculation of various actuarial quantities (such as high percentiles). The bounds for the discounted loss reserve use the maximum likelihood estimates of the parameters in the linear predictor. To incorporate the error arising from the estimation of these parameters Hoedemakers et al. (2004) apply bootstrap methodology. For details on this procedure we refer to the paper. Table 2 compares the Bayesian 95th percentile and the bootstrapped 95th percentile of the lower and upper bound for the different reserves. We bootstrapped 1000 times, computed each time (analytically) the 95th percentile of upper and lower bound and report in Table 2 the 95th percentile of the bootstrap samples obtained in this way.

Table 2
95th Percentile of the Predictive Distribution of the Different Discounted Reserves

Year	Lower Bound	Upper Bound	Bayesian
2	360,725	387,404	436,151
3	700,465	765,451	760,177
4	945,845	982,425	970,535
5	1,441,016	1,513,186	1,448,056
6	1,913,383	1,977,934	1,919,300
7	2,519,292	2,614,564	2,558,208
8	3,557,014	3,702,302	3,641,890
9	4,573,767	4,770,944	4,727,262
10	5,577,925	5,821,804	5,638,301
Overall	20,949,190	21,988,048	20,360,196

Compared with the Bayesian results, Table 2 illustrates that the comonotonicity approach indeed provides relevant information concerning the predictive distribution of discounted loss reserves in a GLM framework.

REFERENCES

DE ALBA, ENRIQUE. 2002. Bayesian Estimation of Outstanding Claims Reserves. *North American Actuarial Journal* 6(4): 1–20.

DHAENE, JAN, MICHEL DENUIT, MARC GOOVAERTS, ROB KAAS, AND DAVID VYNCKE. 2002a. The Concept of Comonotonicity in Actuarial Science and Finance: Theory. *Insurance: Mathematics and Economics* 31(1): 3–33.

———. 2002b. The Concept of Comonotonicity in Actuarial Science and Finance: Applications. *Insurance: Mathematics and Economics* 31(2): 133–61.

HOEDEMAKERS, TOM, JAN BEIRLANT, MARC GOOVAERTS, AND JAN DHAENE. 2003. Confidence Bounds for Discounted Loss Reserves. *Insurance: Mathematics and Economics* 33(2): 297–316.

———. 2005. On the Distribution of Discounted Loss Reserves Using Generalized Linear Models. *Scandinavian Actuarial Journal* (1): 25–45.

KAAS, ROB, JAN DHAENE, AND MARC GOOVAERTS. 2000. Upper and Lower Bounds for Sums of Random Variables. *Insurance: Mathematics and Economics* 27(2): 151–68.

TAYLOR, GREG, AND FRANK ASHE. 1983. Second Moments of Estimates of Outstanding Claims. *Journal of Econometrics* 23: 37–61.

VERRALL, R. J. 2004. A Bayesian Generalized Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving. *North American Actuarial Journal* 8(3): 67–89.

DAVID P. M. SCOLLNIK*

1. INTRODUCTION

I thank Professor Verrall for an interesting paper presenting an innovative and intriguing set of methods for Bayesian reserving. I welcome this opportunity to contribute several comments, some questions for the author, and some of my own suggestions. I will focus my remarks towards the over-dispersed Poisson model introduced by Verrall, leaving it to be understood that similar ones may also be directed towards his negative binomial model.

2. THE OVER-DISPERSED POISSON MODEL

The model for the incremental claims, $C_{i,j}$, described in Section 3 of the paper assumes that $C_{i,j}|\mu_{i,j}, \phi \sim$ independent over-dispersed Poisson, with mean $\mu_{i,j}$. The discussion surrounding equation (3.2) in the paper explains the meaning of the term "over-dispersed" in this context. Specifically, the variable $C_{i,j}/\phi$ has a distribution that is Poisson with a mean of $\mu_{i,j}/\phi$. Hence, conditional on the parameters $\mu_{i,j}$ and ϕ, the quantities $C_{i,j}$ are independent with probability functions given by

$$f(C_{i,j}|\mu_{i,j}, \phi) = \frac{e^{-\mu_{i,j}/\phi}\left(\dfrac{\mu_{i,j}}{\phi}\right)^{C_{i,j}/\phi}}{\left(\dfrac{C_{i,j}}{\phi}\right)!}, \qquad (2.1)$$

for values of $C_{i,j} = 0, \phi, 2\phi, 3\phi, \ldots$.

Now, this over-dispersed Poisson distribution plays two roles in Verrall's model. On the one hand, realizations drawn from the posterior predictive distribution are simulated using (2.1), conditional on the current simulated values of the model parameters $\mu_{i,j}$ and ϕ. By definition, any such simulated values of $C_{i,j}$ are necessarily non-negative. This is the case for predictive draws corresponding to future observations, and also for replicated draws as described below. On the other hand, Verrall adopts a quasi-likelihood approach in order to extend the definition of the model in order to accommodate arbitrary positive or negative observed claims data values.

The method of posterior predictive checking (as described in Gelman, Carlin, Stern, and Rubin [2004, Chapter 6]) is often used in order to determine whether a Bayesian model is consistent with observed data. The idea here is that if "the model fits, then replicated data generated under the model should look similar to observed data. To put it another way, the observed data should look plausible under the posterior predictive distribution" (Gelman et alia, 2004, page 159). To be clear, think of replicated data as the data "that could have been observed, or, to think predictively, as the data we would see tomorrow if the experiment that produced" today's data "were replicated with the same model and the same" parameter values that produced the observed data (Gelman et alia, 2004, page 161).

This leads us to identify a problem inherent in Verrall's Bayesian model, namely, that replicated data values corresponding to negative observed values will never themselves be negative. Professor Verrall notes that "the model can cope with some negative values (as in the data set used in Section 6)" (page 80). However, I believe the point I've identified above recommends against this model's use whenever negative values are contained in the observed data. Otherwise, the posterior predictive distribution will be inconsistent with the observed negative data values. Note that the posterior predictive distribution will also be inconsistent with future expectations if negative values are anticipated in the future data.

I believe it would be useful, and I would personally be grateful, if Professor Verrall were to provide some precise references to papers in the mainstream statistics literature in which an over-dispersed Poisson model has been applied to a data set including negative observed values. I would be interested in reading what the statistics literature says on this subject.

3. THE SCALE PARAMETER ϕ IN THE OVER-DISPERSED POISSON MODEL

Professor Verrall notes that a "fully Bayesian approach would assign a prior distribution for ϕ as well, and integrate it out" (page 70). This is fairly

* David P. M. Scollnik, ASA, PhD, is a Professor in the Department of Mathematics and Statistics, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada, T2N 1N4, e-mail: scollnik@ucalgary.ca.

easy to accomplish within WinBugs, as demonstrated below.

The WinBugs code for the over-dispersed Poisson model given in the Appendix of Verrall's paper includes these lines:

```
zeros[i] ~ dpois(phi[i])
phi[i] ← (mu[i]-Z[i]*log(mu[i])
  +loggam(Z[i]+1))/scale
```

(see page 83). Unfortunately, some of the variable names used in the WinBugs code contained in the Appendix of Verrall's paper differ from the names used in its main body. In the WinBugs code reproduced above, `Z[i]` represents one of the incremental claim amounts $C_{i,j}$ divided by 1000, `mu[i]` is the corresponding mean, and `scale` appears to be the variable that corresponds to the parameter ϕ. These lines of WinBugs code define the contribution made by an observed incremental claim to the likelihood in accordance with (2.1), in the special case that the parameter ϕ is known. Note that the variable `phi[i]` is the minus log likelihood term (up to an additive constant) associated with the particular observed incremental claim, when the parameter ϕ is assumed fixed and known. However, this term is incomplete when ϕ is treated as a stochastic variable. In this case, I believe that the WinBugs code should be changed to:

```
zeros[i] ~ dpois(phi[i])
phi[i] ← (mu[i]-Z[i]
  *log(mu[i]/scale))/scale
  +loggam((Z[i]/scale)+1)
```

The value assigned to `phi[i]` comes from taking minus the logarithm of expression (2.1). Note that the log of the gamma function (i.e., `loggam`) may cause difficulty if the `Z[i]` variable appearing in its argument takes on a negative value. This is another reason for suggesting that Verrall's model is not appropriate when negative values appear in the claims data.

If the modification described above is the only change made to Verrall's code, then the simulation results will be unchanged provided that the variable `scale` (i.e., the parameter ϕ) is held fixed at the same value as before. In his analysis, Verrall uses the value 1.08676 for `scale` and the following assignment appears in the WinBugs code:

```
scale ← 1.08676
```

(see page 84). However, we want to perform the analysis incorporating a prior distribution for the variable `scale` instead of assigning it a fixed value. For the sake of illustration, we assign `scale` a prior distribution that is gamma. Its prior mean is set equal to the previously fixed value of 1.08676, and the analysis will be performed using two different values for its prior variance (i.e., 0.01 and 0.1). To model this in WinBugs, replace the single line of code given above with these:

```
# The prior mean of scale is scale.est,
  the variance is scale.var
scale ~ dgamma(ascale,bscale)
scale.est ← 1.08676
scale.var ← 0.01 # Or use 0.1
ascale ← pow(scale.est,2)/scale.var
bscale ← scale.est/scale.var.
```

The variable `scale` should also be assigned an initial value (e.g., a value around 1.2) when the initial values of the parameters are loaded prior to running this simulation in WinBugs.

I ran Verrall's original model, along with the two models incorporating a stochastic `scale` described above. Each simulation was allowed to burn-in for 50,000 iterations, and then allowed to run for another 10,000 iterations in order to obtain the results shown below. Table 1 shows the results of running Verrall's original WinBugs code for the over-dispersed Poisson model. The Bayesian Mean Reserve (BMR) and the Bayesian Pre-

Table 1

Over-Dispersed Poisson Model: Mean and Prediction Error of Reserves: Verrall's Base Model with Vague Priors and Fixed ϕ

	Bayesian Mean Reserve	Bayesian Prediction Error	Bayesian Prediction Error %
Year 2	190	688	362%
Year 3	691	1,289	187
Year 4	1,743	1,986	114
Year 5	2,901	2,486	86
Year 6	3,769	2,596	69
Year 7	5,569	3,329	60
Year 8	11,090	5,216	47
Year 9	10,870	6,369	59
Year 10	17,240	14,240	83
Overall	54,070	19,840	37

diction Error (BPE) are the estimated mean and standard deviation of the posterior distribution as returned by WinBugs. The Bayesian Prediction Error % is the ratio of the BPE to the BMR (expressed as a percentage). These results are essentially the same as those reported in Table 2 of Verrall's own paper (see page 77). Tables 2 and 3 show the results for the model when the variable `scale` (i.e., the parameter ϕ) is treated as a stochastic variable with a prior distribution that is gamma, with a mean of 1.08676 and a variance of 0.01 and 0.1, respectively. It is apparent that the reserves in Tables 1 and 2 are smaller and much less variable than the ones reported in Table 3. Note that the 95% posterior probability intervals reported by WinBugs for the overall reserve are (25,000, 102,200), (21,200, 115,200), and (9,069, 189,300), for the models underlying Tables 1, 2, and 3, respectively. Also note that for the model underlying Table 2, the parameter ϕ has an estimated posterior mean and standard deviation of 1.501 and 0.1122, respectively, as returned by WinBugs. For the model underlying Table 3, these values are 4.197 and 0.569.

The DIC Tool dialog box in WinBugs may be used to monitor the Deviance Information Criterion (DIC) (Spiegelhalter, Best, Carlin, and van der Linde, 2002) and related statistics. These statistics can be used to assess model complexity and compare different models. The model with the smallest DIC is estimated to be the model that would best predict a replicate dataset (as discussed above) of the same structure as that currently observed. I monitored the DIC for each of the three models I simulated. For the model underlying Table 1 (with `scale` fixed and set equal to 1.08676), the DIC was 213.4; for the model underlying Table 2 (with a stochastic variable `scale` with prior variance of 0.01), it was 190.4; and for the model underlying Table 3 (with a stochastic variable `scale` with prior variance of 0.1), it was 133.8. These results suggest that the third model is the best of the three.

This example shows that it is relatively straightforward to implement a fully Bayesian analysis of Verrall's over-dispersed Poisson model, incorporating a prior distribution for ϕ, in WinBugs. In this case, the result was a better fitting model. Furthermore, ignoring the stochastic variability in ϕ resulted in an underestimation of the various years' reserves, and in a significant understatement of their variability. Verrall's negative binomial model may also benefit from the introduction of a stochastic, rather than fixed, parameter ϕ. In either case, the prior distribution for ϕ is probably best determined on the basis of expert opinion or on the results of a prior study. If a vague, by which we mean improper or proper but relatively widely dispersed, prior distribution is used for ϕ, then it may be necessary to tighten up some of the other parameters' prior distributions in order for WinBugs to successfully compile and run.

Table 2

Over-Dispersed Poisson Model: Mean and Prediction Error of Reserves: Stochastic ϕ with Its Prior Variance Equal to 0.01

	Bayesian Mean Reserve	Bayesian Prediction Error	Bayesian Prediction Error%
Year 2	169	806	477%
Year 3	636	1,405	221
Year 4	1,698	2,252	133
Year 5	2,841	2,804	99
Year 6	3,757	3,032	81
Year 7	5,556	3,847	69
Year 8	11,110	6,166	55
Year 9	11,150	7,803	70
Year 10	18,210	18,770	103
Overall	55,130	25,000	45

Table 3

Over-Dispersed Poisson Model: Mean and Prediction Error of Reserves: Stochastic ϕ with Its Prior Variance Equal to 0.1

	Bayesian Mean Reserve	Bayesian Prediction Error	Bayesian Prediction Error%
Year 2	182	1,282	704%
Year 3	789	2,913	369
Year 4	1,960	4,370	223
Year 5	3,100	5,068	163
Year 6	4,051	5,473	135
Year 7	5,742	6,866	120
Year 8	11,650	11,050	95
Year 9	11,710	13,910	119
Year 10	20,440	35,170	172
Overall	59,620	46,370	78

ACKNOWLEDGEMENTS

The author acknowledges support from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

GELMAN, ANDREW, JOHN B. CARLIN, HAL S. STERN, AND DONALD B. RUBIN. 2004. *Bayesian Data Analysis*. 2nd ed. Boca Raton, United States: Chapman & Hall/CRC.

SPIEGELHALTER, DAVID J., NICOLA G. BEST, BRADLEY P. CARLIN, AND ANGELIKA VAN DER LINDE. 2002. Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society, Series B: Statistical Methodology* 64: 583–616.

AUTHOR'S REPLY

I would like to thank the discussants for their interesting contributions. As both point out, Bayesian statistics offer great possibilities for the actuary and we have some good suggestions here for what could be useful. For readers who would like to try the Bayesian methods using the software WinBUGS, I have included more detailed instructions in the Forum of the Casualty Actuarial Society, Verrall (2004). This paper also contains some other examples of how Bayesian methods might be useful, which are covered in more detail in Verrall and England (2005). This is clearly an expanding area, as is shown by the contributions of the discussants and the papers mentioned above. I expect that it will continue to develop fast, and that the presently predominant methods will be replaced by better approaches in the future.

Professor Scollnik makes some criticisms of the paper, which could be argued arise because of my desire to provide stochastic methods that stay close to the present actuarial procedures. Given complete freedom to decide on what data should be supplied, what methods are acceptable to management, regulators, etc., and complete freedom to use methods as complicated and sophisticated as might be desirable, I would probably do things differently. Faced with the reality of the pressures of commercial life, my desire has been to show how it is possible to move gradually towards a stochastic approach. Once this has been achieved (and it is beginning to happen), there will be more possibilities to use other methods, and I am sure that researchers, such as Scollnik, will be able to provide better models. For example, Scollnik helpfully shows how the estimation of the dispersion parameter may be included in the Bayesian model. This is something I have also looked at, using the suggestions of West (1985), but I decided that it would be better to make a consistent comparison with the non-Bayesian methods where the usual procedure is to ignore the effect of the estimation of the dispersion parameter.

Strictly, the use of the over-dispersed Poisson or negative binomial implies that the data consists of multiples of the dispersion parameter. In practice, this can be over-looked for the observed past data by using the "zeros trick," as I did in the WinBUGS code here. It is obviously more problematic when simulating future observations: The properties of the predictive distributions will be consistent with the assumptions made, but the distribution of a single cell in the run-off triangle will appear unrealistic. Another approach is to use an alternative forecast distribution—such as the Gamma distribution—parameterised such that the mean and variance equal their theoretical values. In making this pragmatic compromise, it could be argued that we are introducing a conceptual inconsistency. An alternative to the over-dispersed negative binomial distribution is to use the normal distribution. This was applied in England and Verrall (2002) in the context of the chain-ladder technique, and an equivalent approach could also be used for the Bornhuetter-Ferguson technique, and it is possible that a strict Bayesian would prefer this.

REFERENCES

VERRALL, R. J. 2004. *Obtaining Predictive Distributions for Reserves Which Incorporate Expert Opinion.* Casualty Actuarial Society Forum, Fall.

VERRALL, R. J., AND P. D. ENGLAND. 2005. Incorporating Expert Opinion into a Stochastic Model for the Chain-Ladder Technique. *Insurance: Mathematics and Economics.* Forthcoming.

WEST, M. 1985. Generalised Linear Models Scale Parameters, Outlier Accommodation and Prior Distributions. In *Bayesian Statistics 2,* ed. A. P. Dawid, D. Heckerman, A. F. M. Smith and M. West. Amsterdam: North-Holland.

"A FRAMEWORK FOR LONG-TERM ACTUARIAL PROJECTIONS OF HEALTH CARE COSTS: THE IMPORTANCE OF POPULATION AGING AND OTHER FACTORS," HOWARD J. BOLNICK, OCTOBER 2004

JOHN A. BEEKMAN*

Howard Bolnick has provided readers with an excellent comprehensive study of key health care cost drivers, three plausible futures for health care, and their potential costs. This paper will help actuaries and health economists in their stewardship of health care systems.

The paper refers to Fries (1980), which determined upper limits to natural life expectancy and life span and demonstrated a rectangularization of the survival curve, assuming a natural genetic limit to life. This reference occurs in Sections 5.1 (including Fig. 11) and 5.2. The discussant found very useful another paper in *The New England Journal of Medicine,* namely, Katz et al. (1983), which refers to the concept of "active life expectancy." The discussant wrote a lengthy paper on that concept (Beekman 1988) and how it could be used by actuaries. To quote from its abstract:

> Active life expectancy provides a measure of the expected number of years of independence in certain daily living activities. . . . The paper presents mathematical expressions for active life expectancies, projections of active life expectancies through the year 2040, and some material on the estimation of long-term health care costs. The last section presents a formula for a net premium for long-term health care for the elderly, using some of the mathematical foundations of active life expectancies. The formula is applied in an example which uses Medicare and medical data.

I will use a portion of Table 3 in Beekman (1988, p. 16) to create Table 1. The e_{65}^{0} values come from the Alternative II columns in Table 10 of Wade (1986). The notation (ae_{65}^{0}) refers to the complete version of active life expectancy at age 65.

The later publication Beekman and Frye (1991) projected active life expectancies to the year 2080, under three sets of assumptions, for 65-year-old men and women. The Alternative I, II, and III sets of assumptions use low, intermediate, and high values for e_{65}^{0} and come from Table 10 of Wade (1989). A portion of Table 7 from Beekman and Frye (1991) is included here as Table 2.

Table 1
Projections of Active Life Expectancies

Year	Male		Female	
	e_{65}^{0}	(ae_{65}^{0})	e_{65}^{0}	(ae_{65}^{0})
2010	16.0	11.4	20.8	11.2
2020	16.3	11.6	21.3	11.5
2040	17.1	12.1	22.2	12.0

Table 9 of Beekman and Frye (1991) provided estimates of the expected numbers of years of dependency in activities of daily living (ADL) prior to death of 65-year-old men and women, under the same Alternative I, II, and III sets of assumptions. A portion of that table is included here as Table 3.

This table is relevant to footnote 7 of Bolnick (2004, p. 23), in which disability adjusted life years (DALYs) are described as "the years of life expectancy lived in a state of less than good health; it is the difference between life expectancy and healthy [adjusted] life expectancy (HALE), where HALE is that portion of life expectancy lived in a state of good health. See the World Health Report (2002) for a complete description of DALY and HALE and how they are calculated."

Section 5.3.4 of Bolnick (2004, pp. 14–15), including Figure 12, "compares healthy life expectancy to the cost of health care in 2001 for each of the 19 countries included in this study." The paper also refers (pp. 21–22) to Jacobzone (1999), which compares severe-disability-free life expectancy to life expectancy, both age 65, for three countries (United States, Japan, France) for 1970–90 and four countries (Norway, United Kingdom, Australia, Canada) for 1976–91. These comparisons appear in Figure 14 Bolnick (2004, p. 22).

Because of the discussed paper's references to "healthy life expectancy," "severe-disability-free

* John A. Beekman is Lincoln National Corp. Professor Emeritus of Actuarial Science, Ball State University, Department of Mathematical Sciences, Muncie, IN 47306-0490, 00jabeekman@bsu.edu.

Table 2
Projections of Life and Active Life Expectancies

Year	Alternative I				Alternative II				Alternative III			
	Male		Female		Male		Female		Male		Female	
	e_{65}^{0}	(ae_{65}^{0})	e_{65}^{0}	(ae_{65}^{0})	e_{65}^{0}	(ae_{65}^{0})	e_{65}^{0}	(ae_{65}^{0})	e_{65}^{0}	(ae_{65}^{0})	e_{65}^{0}	(ae_{65}^{0})
2010	15.2	13.83	19.0	14.63	16.0	14.56	20.1	15.48	17.0	15.47	21.2	16.32
2020	15.3	13.92	19.2	14.78	16.4	14.92	20.5	15.79	17.8	16.20	22.0	16.94
2030	15.5	14.11	19.4	14.94	16.8	15.29	20.9	16.09	18.6	16.93	22.9	17.63
2040	15.7	14.29	19.6	15.09	17.1	15.56	21.4	16.48	19.3	17.56	23.7	18.25
2050	15.9	14.47	19.8	15.25	17.5	15.93	21.8	16.79	20.1	18.29	24.5	18.87
2060	16.0	14.56	20.0	15.40	17.8	16.20	22.2	17.09	20.9	19.02	25.4	19.56
2070	16.2	14.74	20.2	15.55	18.2	16.56	22.6	17.40	21.7	19.75	26.2	20.17
2080	16.3	14.83	20.3	15.63	18.5	16.84	23.0	17.71	22.5	20.48	27.0	20.79

life expectance,'' and DALYs, it seemed appropriate to direct the readers to the projections of ''active life expectancy'' and ''expected number of years of dependency in ADL'' contained in Beekman (1988) and Beekman and Frye (1991).

Howard Bolnick's excellent paper is a major reference for actuaries and health economists in their careful projections of health care costs for future years.

Table 3
Expected Number of Years of Dependency in ADL: $e_{65}^{0} - (ae_{65}^{0})$

Year	Alternative I		Alternative II		Alternative III	
	Male	Female	Male	Female	Male	Female
2010	1.37	4.37	1.44	4.62	1.53	4.88
2020	1.38	4.42	1.48	4.71	1.60	5.06
2030	1.39	4.46	1.51	4.81	1.67	5.27
2040	1.41	4.51	1.54	4.92	1.74	5.45
2050	1.43	4.55	1.57	5.01	1.81	5.63
2060	1.44	4.60	1.60	5.11	1.88	5.84
2070	1.46	4.65	1.64	5.20	1.95	6.03
2080	1.47	4.67	1.66	5.29	2.02	6.21

REFERENCES

BEEKMAN, JOHN A. 1988. Active Life Expectancies and Their Actuarial Uses. Technical Report no. 79. Ball State University, Department of Mathematical Sciences.

BEEKMAN, JOHN A., AND WILLIAM B. FRYE, 1991. Projections of Active Life Expectancies. *Actuarial Research Clearing House* 2: 73–87.

FRIES, JAMES F. 1980. Aging, Natural Death, and the Compression of Morbidity. *New England Journal of Medicine* 303(3): 130–35.

JACOBZONE, STEPHANE, 1999. Aging and Care for Frail Elderly Persons: An Overview of International Perspectives. OECD Labor Market and Social Policy Occasional Paper no. 38. April.

KATZ, S., L. G. BRANCH, M. H. BRANSON, J. A. PAPSIDERO, J. C. BECK, AND D. S. GREER. 1983. Active Life Expectancy. *New England Journal of Medicine* 309: 1218–24.

WADE, ALICE H. 1986. *Social Security Area Population Projections, 1986.* Actuarial Study no. 97. Baltimore, Md: Office of the Actuary, Social Security Administration.

———. 1989. *Social Security Area Population Projections, 1989.* Actuarial Study no. 105. Baltimore, Md: Office of the Actuary, Social Security Administration.

WORLD HEALTH REPORT. 2002. *Reducing Risks, Promoting Healthy Life.* Geneva: World Health Organization.