ISSN 0709-9231

Cahier 8838

APPLICATIONS OF THE GB2 FAMILY OF DISTRIBUTIONS IN MODELING INSURANCE LOSS PROCESSES[1]

by

J. David Cummins*
Georges Dionne**
James B. McDonald***
and
B. Michael Pritchett***

* Department of Insurance, Wharton School, University of Pennsylvania, Philadelphia, PA 19104, U.S.A.

** Département de Sciences économiques et Centre de recherche sur les transports, Université de Montréal, C.P. 6128, Succursale A, Montréal (Québec) CANADA H3C 3J7.

*** Department of Managerial Economics, Brigham Young University, Provo, UT 84602, U.S.A.

November 1988

[1] The Social Sciences and Humanities Research Council of Canada and the Huebner Foundation provided financial support to Georges Dionne. Charles Vanasse was very helpful in the estimation of the parameters made in Montreal.

Ce cahier est aussi publié par :

Centre de recherche sur les transports - Publication #602

ABSTRACT

This paper investigates the use of a four-parameter family of probability distributions, the generalized beta of the second kind (GB2), for modeling insurance loss processes. The GB2 family includes many commonly used distributions such as the log-normal, gamma and Weibull. The GB2 also includes the Burr and generalized gamma distributions and has significant potential for improving the distributional fit in many applications involving thin or heavy-tailed distributions. Members of the GB2 family can be generated as mixtures of well-known distributions and provide a model for heterogeneity in claims distributions. Examples are presented which consider models of the distribution of individual and of aggregate losses. The results suggest that seemingly slight differences in modeling the tails can result in large differences in reinsurance premiums and quantiles for the distribution of total insurance losses.

Key words : generalized beta of the first kind; generalized beta of the second kind; log t; generalized gamma; Pearson; Burr; Kappa; Pareto; Lomax; Log Cauchy; lognormal; beta; gamma; Weibull, Fisk; Rayleigh; uniform; exponential; inverse distributions; reinsurance.

\- - - - - -

RÉSUMÉ

Ce texte suggère l'utilisation de la famille des distributions de probabilités à quatre paramètres, la beta généralisée du deuxième type (GB2), pour analyser les pertes d'assurance. La GB2 a comme cas particulier les distributions log-normale, gamma et Weibull. Elle comprend également comme cas particuliers la Burr et la gamma généralisée et elle est très appropriée pour améliorer l'estimation des distributions ayant des queues très épaisses. La plupart des membres de la famille de distribution proposée peuvent être obtenus par des combinaisons de distributions bien connues. Les résultats démontrent que des faibles différences dans la modélisation des queues peuvent générer des grandes différences dans le calcul des primes de réassurance et dans les quantiles des distributions des pertes assurées totales.

Mots-clés : beta généralisée du premier et du second type; log t; gamma généralisée; Pearson; Burr; Kappa; Pareto; Lomax; log Cauchy; log-normale; beta; gamma; Wiebull; Fisk; Rayleigh; uniforme; exponentielle; distributions inverses; réassurance.

APPLICATIONS OF THE GB2 DISTRIBUTION IN MODELING INSURANCE LOSS PROCESSES

1. Introduction

One of the classic problems in collective risk theory is the estimation of the distribution of total annual claims, F(X),

$$X = \sum_{i=1}^{n} x_i$$

where n denotes the number of claims in a time period and xi is the amount (severity) of the ith claim. Thus, estimation of total claims usually involves fitting the frequency and severity distributions, testing-for-goodness of fit, and then combining the distributions to yield F(X). The model of F(X) is then used to estimate premiums, risk loadings, reinsurance premiums, and other decision variables such as the maximum probable yearly aggregate loss (MPY).

Distributions used for frequency include the Poisson, the negative binomial and the logarithmic series distribution (see, for example, Ferreira [1970], Seal [1969], and Cummins and Wiltbank [1983]). Among the distributions that have been considered for severity are the exponential, gamma, loggamma, lognormal, Pareto, and log-t, (Beard, Pentikainen and Pesonen [1969], Benckert [1962], Cummins and Freifelder [1978], Cummins and Wiltbank [1983], Hewitt [1970], Mandlebrot [1964], Paulson [1984], Seal [1969], and Shpilberg [1977]).

Traditionally, calculating F(X) directly was considered a difficult problem, and various approximation formulas, such as the normal-power and gamma, received considerable attention (see, for example, Beard, Pentikainen, and Pesonen [1984]). In recent years, developments in risk theory and advances in computing have made virtually exact calculation of F(X) quite feasible. The most prominent approaches are simulation (e.g., Roy and Cummins [1985], fast Fourier transforms (FFTs) (Paulson [1984]), and the Panjer recursion algorithm (Panjer [1981]).

Paralleling the developments in the calculation of F(X) have been more sophisticated approaches to the estimation of the frequency and severity distributions. The traditional approach was the use of tractable one or two-parameter distributions such as those mentioned above. In many cases, these distributions were selected not necessarily because they were in some sense "best" but because of the feasibility of estimating parameters and computing fractiles. Because insurance claims distributions are often heavy-tailed, restricting the set of candidate distributions can lead to serious underestimation of tail fractiles reinsurance premiums, and other variables (Cummins and Freifelder [1978]).

Recent advances have opened up a much wider range of probability distributions for use in modeling insurance claims processes. Hogg and Klugman

[1984] discuss many alternative models for loss distributions as well as related issues of estimation and inference. Paulson has made extensive use of the stable family of distributions, which includes the one-tail Pareto, the normal, and the Cauchy distributions, among others, as special cases (e.g., Paulson and Faris [1985]). Aiuppa [1986], utilizes a computer program that can estimate parameters and compute percentile points of the distribution functions for any member of the Pearson family.[1] McDonald [1984], considers generalizations of the beta distribution of the first type (Pearson Type I) and of the second type (Pearson Type VI), which will be denoted GB1 and GB2, respectively. Venter [1984] introduced the GB2 in the actuarial literature as the transformed beta. Applications of the GB1 and GB2 in the economics literature are described in McDonald [1984], McDonald and Butler [1987], and McDonald and Richards [1984].

The purpose of this article is to investigate the use of the GB2 family as a model for continuous distributions in non-life insurance. Thus, the GB2 is proposed as a potential model for both aggregate losses and loss severity. The GB2 provides an extremely flexible functional form that can be used to model highly skewed loss distributions such as those typically observed in non-life insurance. The use of the GB2 is illustrated below by modeling the fire losses of a major university, using the Cummins and Freifelder [1978], data set.

In working with the GB2, our approach is to begin with the simplest (one or two-parameter) applicable members of the family and to move up the distribution tree to the most general (four parameter) GB2 distribution. Improvements in goodness-of-fit obtained by moving to higher levels of the tree are measured and discussed. The GB2 family can be used either with the untransformed data or with natural logs of the data. The log-GB2 may be better in some instances if the data are characterized by extremely heavy tails. For our data set, the log transformation proved to be unnecessary.

Our model of the aggregate claims distribution is based on the moments of a time series of observed total losses. Simulation is also used to obtain F(X) from the underlying frequency and best-fitting severity distributions. We do not formally consider other methods of compounding frequency and severity to yield F(X). However, GB2 severity distributions could easily be incorporated into the Panjer and FFT algorithms.

The paper is organized as follows: The generalized distributions are discussed in section 2. Two applications of these distributions are then considered: estimating the distribution of aggregate losses in section 3, and estimating the severity distribution in section 4. Severity distributions are estimated with both grouped and ungrouped data to explore the issue of accuracy loss from using grouped data. Section 5 discusses the impact of model selection and estimation technique on the severity fractiles, reinsurance premiums, and simulated total claims distributions. Section 6 concludes the paper.

2. The Generalized Distributions

It is useful to define three very flexible distributions: the generalized gamma (GG), the generalized beta of the first kind (GB1) and the generalized beta of the second kind (GB :

$$(1)\quad GG(x;a,b,p) = \frac{|a|x^{ap-1}\, e^{-(x/\beta)^a}}{\beta^{ap}\,\Gamma(p)} \qquad 0 \le x$$
$$= 0 \qquad \text{otherwise}$$

$$(2)\quad GB1(x;a,b,p,q) = \frac{|a|x^{ap-1}\,(1 - (x/b)^a)^{q-1}}{b^{ap}\,B(p,q)} \qquad 0 \le x^a \le b^a$$
$$= 0 \qquad \text{otherwise}$$

$$(3)\quad GB2(x;a,b,p,q) = \frac{|a|x^{ap-1}}{b^{ap}\,B(p,q)(1 + (x/b)^a)^{p+q}} \qquad 0 \le x$$
$$= 0 \qquad \text{otherwise}$$

The generalized gamma is a three parameter distribution and is a limiting case of both the four parameter GB1 and GB2 distributions. The parameters in these distributions determine the shape and location of the density. The parameter "b" is a scale factor; "b" is also an upper (lower) bound for GB1 variables as the parameter "a" is positive (negative). Unlike the GB1 or the beta of the first kind (B1), which is mentioned in risk theory texts (for B1 see, for example, Buhlmann [1970]), the GB2 has no upper limit and hence is likely to be applicable for severity distributions and other risk theory applications where the upper tail has no theoretical boundary. The GB1 and GG have defined moments up to order h where $h/a > -p$. For "a" positive, moments of all positive integer orders are defined. The GB2 has integer moments of order up to h where $-p < h/a < q$.

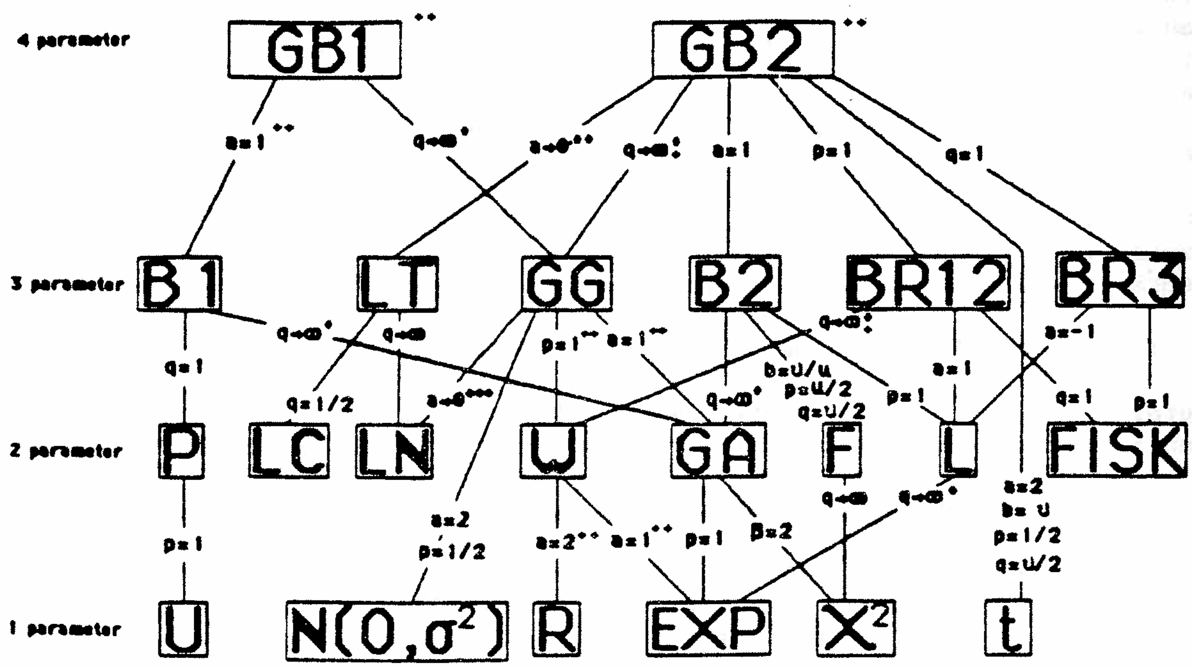
Thus, the GB2 provides models for distributions characterized by thick tails. The relationship between the density and parameters is complex, but generally speaking the larger the values of the parameters "a" or "q" the thinner the tails of the density function. The relative values of "p" and "q" are important in determining the skewness of the distribution and the GB2 permits positive as well as negative skewness.

The moments and expressions for the distributions of the GG, GB1 and GB2 are given in Table 1. The terms ${}_2F_1[\ \ ]$ and ${}_1F_1[\ \ ]$, respectively, denote the hypergeometric and confluent hypergeometric distributions and, except for special cases, involve infinite series (Abramowitz and Stegun [1965]). However, one very flexible member of the GB2 family, the Burr 12 (GB2 with $p = 1$) has a closed form distribution function; and the fractiles of the four-parameter GB2 can be obtained by numerical integration.

Table 1
Distribution and Moments

Model	Distribution function	Moments
GG	$\frac{(x/\beta)^{ap}\exp(-(x/\beta)^{a})}{\Gamma(p+1)} \; {}_1F_1 \left[1, p+1; (x/\beta)^{a} \right]$	$\frac{b^{h}\Gamma(p+h/a)}{\Gamma(p)}$
GB1	$\frac{(x/h)^{ap}}{pB(p,q)} \; {}_2F_1 \left[p, 1-q; p+1; (x/b)^{a} \right]$	$\frac{b^{h}B(p+h/a,q)}{B(p,q)}$
GB2	$\frac{\left[\left(\frac{(x/b)^{a}}{1+(x/b)^{a}} \right) \right]^{p}}{pB(p,q)} \; {}_2F_1 \left[p, 1-q; p+1; \frac{(x/b)^{a}}{1+(x/b)^{a}} \right]$	$\frac{b^{h}B(p+h/a, q-h/a)}{B(p,q)}$

Figure 1
Distribution Family Tree



* $q \to \infty$ with $b = q^{1/a}\beta$
** $a \to 0$ with $b = (q\sigma^2 a^2)^{1/a}$, $p = (au+1)/\sigma^2 a^2$
*** $a \to 0$ with $b = (\sigma^2 a^2)^{1/a}$, $p = (au+1)/\sigma^2 a^2$

* If $p \to \infty$ with $b = p^{1/a}\beta$, the inverse form of the distribution is obtained.
** The inverse form is obtained if the sign of a is changed.

The flexibility of these distributions can be illustrated using figure 1 (from McDonald and Richards [1987] or McDonald and Butler [1987]). The GB2 is seen to include the log-t (LT), generalized gamma (GG), beta of the second kind (B2), Burr types 3 and 12 (BR3 and BR12), log-Cauchy (LC), lognormal (LN), gamma (GA), Weibull (W), Lomax (L), Fisk, Rayleigh (R), and exponential (EXP) as special or limiting cases. The GB2 family can also be defined for the natural logs of the variable of interest, yielding the log-GB2. This transformation may be useful for very heavy tailed distributions, providing the restriction that $\log(x) > 0$ does not create problems. The GB1 includes the beta of the first kind (B1), power (P), uniform (U) and generalized gamma with related distributions appearing as special or limiting cases.

When the parameter "a" is negative, (1), (2) and (3) admit inverse distributions. These are distributions of the random variable y, obtained by making the reciprocal transformation $x = 1/y$. Examples are the inverse Burr and the inverse gamma, which includes the Pearson type V as a special case as well as inverse functions of other special and limiting cases related to the generalized gamma. Thus the GB1 and GB2 include many members of the Pearson family of distributions (Ord [1972] and Elderton [1964]). Other models included in the GB2 family, which are not Pearson distributions, include the Burr types 3 and 12, Fisk, Weibull, generalized gamma and log-t. Additional details are included in McDonald [1984], McDonald and Butler [1987], McDonald and Richards [1987] or Venter [1984].

Each of these distributions may be well suited for a particular type of data and not for another. For example, the exponential distribution often is not applicable to insurance claims distributions because its mode is zero and it imposes a restrictive relationship between the mean (equal to the distribution's one parameter, b) and the variance (b^2). Two-parameter models allow increased flexibility in not only modeling the mean but possibly the variance and then implying particular corresponding values for the skewness and kurtosis. The four parameter models can allow for independent adjustments to the skewness and kurtosis as well as the mean and variance.

Heterogeneity is a problem which is often encountered in insurance data. Heterogeneity frequently results in distributions with thick tails. Hogg and Klugman [1983] indicate how mixture distributions provide an approach to modeling unobservable heterogeneity. The GB2 provides a mixture interpretation which allows, but does not require heterogeneity. The GB2 can be shown to arise from a structural distribution which is $GG(x;a,\theta,p)$ where the scale parameter θ is distributed as a $GG(\theta;-a,b,q)$. The limiting case of large values of q corresponds to homogeneity.

Each special case of the GB2 can be interpreted as a mixture. Some important cases (LT, BR3, BR12, B2 and Lomax) are summarized in Table 2. For example, the log-t has been shown to be a lognormal mixed with an inverse gamma (Cummins and Freifelder [1978] and Hogg and Klugman [1984]).[2] Heterogeneity can be tested by estimating the models in the first two columns of Table 2 using maximum likelihood estimation and testing for significant differences using a likelihood ratio test. See McDonald and Butler [1987] and Venter [1984] for more details. Thus, the GB2 has a theoretical justification as a representation of claims arising from a heterogeneous population of exposures.

The distribution functions can be "evaluated" using numerical integration or equivalent series representations. Either approach can be implemented on a personal computer, given appropriate software. The distribution for the GB2 can be obtained from the F distribution by utilizing the transformation: $y = (q/p)(x/b)^a$. If x is GB2, then y will be F with degrees of freedom $d_1 = 2p$ and $d_2 = 2q$.

Table 2
Some Mixture Distributions

Observed Distribution	Structural Distribution	Parameter Mixing Distribution
$GB2(x;a,b,p,q)$	$GG(x;a,\theta,p)$	$GG(\theta;-a,b,q)$
$LT(x;u,\sigma^2,q)$	$LN(x;u,\theta)$	$GG(\theta;a=-1,\sigma^2q,q)$
$BR3(x;a,b,p)$	$GG(x;a,\theta,p)$	$W(\theta;-a,b)$
$BR12(x;a,b,q)$	$W(x;a,\theta)$	$GG(\theta;-a,b,q)$
$B2(x;b,p,q)$	$GA(x;\theta,p)$	$GG(\theta;a=-1,b,q)$
$L(x;b,q)$	$EXP(x;\theta)$	$GG(\theta;a=-1,b,q)$

The importance of the GB2 distribution for risk theory is that it has great flexibility due to the availability of four parameters. In addition, it encompasses many of the traditional distributions as special cases. For riskier loss processes another set of distributions is provided by the log GB2.

3. Aggregate Losses: The GB2 and MPY

The aggregate loss function descriptive of a portfolio of insurance policies was defined in section 1. Computation of the Maximum Probable Yearly Aggregate Loss (MPY) from the estimated distribution of X illustrates one use of the generalized distributions. The MPY is identified by a value of X in the upper tail of the distribution of total annual losses that will be equaled or exceeded with a probability no greater than a specified value.

Several approaches to modeling the distribution of X have been considered (see Cummins and Freifelder [1978]). As mentioned above, one method of obtaining the aggregate loss distributions is by estimating and then compounding frequency and severity distributions. One reason for decomposing the problem in this way has to do with sample size. One year's observations on a relatively large pool may prove sufficient to estimate frequency and severity distributions with a high degree of confidence, providing that data have been maintained in the appropriate degree of detail. By contrast, the aggregate losses of the pool for that year constitute only one observation from F(X).

If the pool is relatively small and/or if frequency and severity data have not been maintained in a sufficient degree of detail, it is sometimes possible to estimate F(X) from a time series of observations on total claims, X.[3] If a sufficient number of observations on X is available, it may be advisable to fit the GB2 distribution to this sample even if frequency and severity data have also been maintained. In this instance, the GB2 distribution fitted to the aggregate data provides an alternative to the distribution obtained by compounding frequency and severity. Thus, fitting the GB2 to the aggregate data will be useful, at a minimum, to check the reasonableness and stability of the estimates obtained from the alternative methods.

The method of choice for estimating the GB2 from aggregate data is maximum likelihood. This method can be applied either to individual observations or to grouped data. In either case, N must at least exceed the number of parameters, and a much larger sample is desirable. Maximum likelihood estimates can be obtained by maximizing the usual loglikelihood function for a sample consisting of individual observations on X or, for grouped data, the loglikelihood function of the corresponding multinomial distribution.

The method of moments offers the intriguing possibility of using the GB2 to model total claims when separate frequency and severity distributions are available and when an adequate number of observations on X is not available. In this regard, the method-of-moments GB2 would be an alternative to an approximation method such as the normal power or to a compounding approach. The prospects for this type of application are somewhat limited, however, due to the fact that the moments of the GB2 distribution are not always defined. (For one member of the family, the log-t, none of the moments are defined.) The statistical efficiency of this method is questionable if the data are from a population without the requisite number of theoretical moments (Ord [1972]).[4]

Method of moments estimators are obtained by solving the system of four nonlinear equations obtained by setting the first four theoretical moments of the GB2 equal to the corresponding empirical moments.[5] The moments of X can be computed directly from the sample or inferred from severity and frequency data.[6]

In this section we analyze the data on aggregate fire loss at a major university reported in Cummins and Freifelder [1978]. These data are reproduced in Table 3. The GB2, LT, GG, B2, BR3, BR12, LN, W, GA, Lomax and exponential are fit to the data in Table 3 using maximum likelihood estimation. The estimated parameter values and corresponding loglikelihood values are given in Table 4.[7] The estimates in Table 4 are obtained directly from the data on aggregate losses.

Under the usual procedures for analyzing loss distributions (see, for example, Hogg and Klugman [1984]), one would not bother to estimate the exponential for most claims distributions, due to its shape and restrictiveness with regard to moments. The reason for estimating the exponential distribution when using the GB2 family is to facilitate comparisons of the alternative members of the GB2 family.

Table 3
Fire Loss Experience of a Major University

Year	Total Losses	No. of Exposure Units
1950	71280	270
1951	3671	273
1952	18664	276
1953	8784	279
1954	3966	282
1955	30892	287
1956	631626	292
1957	11464	297
1958	127194	302
1969	4950	308
1960	30452	314
1961	8028	312
1962	14790	310
1963	9480	308
1964	8676	306
1965	114198	305
1966	5150	310
1967	105864	315
1968	32814	320
1969	41340	325
1970	46284	329
1971	12230	330
1972	19418	322
Average	59183	303

mean = 59183 skewness = 3.94
Var = 1.61878×10^{11} kurtosis = 17.84

The likelihood ratio statistic can be used to test for statistically significant differences between "nested" models. For such cases, twice the difference between the loglikelihood values is asymptotically distributed as a Chi-square with degrees of freedom equal to the difference between the number of free parameters in the two models being compared. Based on the log-likelihood values, several of the fitted distributions appear to be "observationally" equivalent. In particular, the GB2, GG, B2 and Burr 3 distributions fit the data equally well. Other distributions could not be rejected as population distributions, e.g. the LT, GG, Burr 12, Lognormal, and Lomax distributions. However, 23 observations could hardly be expected to validate the use of asymptotic tests.

Table 5 reports the MPY estimates (quantiles) based on the estimated models reported in Table 4. The estimates for the .01 level reported by Cummins and Freifelder using normal approximations, Chebyshev, and normal power methodologies are also included in Table 5.

The alternative models provide very different results for the MPY. Given the great diversity in the estimates of MPY, there appears to be an open question as to the validity of some estimators. The Monte Carlo approach also appears to be very sensitive to the assumed distributions. These results reinforce the importance of using appropriate distributions.

In order to investigate this question further in another environment, one in which the actual MPY is known, we assumed that annual losses were distributed according to a known distribution. A random sample of total losses was generated from a known hypothetical population and alternative estimators of MPY obtained and compared with the "true" MPY.

The "true" distribution for total losses was assumed to be defined by:

$$GB2(L; a=1.25,\ b=10,\ p=100,\ q=5).$$

One hundred observations for total losses were generated from this population using a random number generator. The sample skewness and kurtosis were 1.93 and 8.74 respectively and do not depart from normality as much as the fire loss example considered previously. The GB2, B2, LT, and LN were fit to the data using maximum likelihood estimation. The generated data were such that the differences of fit among some alternative distributions were not statistically significant. Of course, the parameters of GB2 could have been selected to yield data in which the differences were significant. Quantiles were calculated corresponding to each of these estimated distributions along with three approximation method. The results are summarized in Table 6.

Some of the results of this example are as expected. The GB2 and normal power provide the best estimates of the quantiles (.01) and appear in this example to have thicker tails than either the LT or LN. The normal power results are expected, given that skewness is less than 2.[8] The Chebyshev results are consistent with the comments of Cummins and Freifelder [1978], that the high MPY's computed by this method tend to imply very conservative (high) premium structures.

Table 4
Parameter Estimates
Unadjusted Data

Model	Parameter Estimates				Loglikelihood
	$a(\mu)$	$b(\sigma)$	p	q(d.f.)	LL
GB2	1.2688	4.3336	14078.1	.68389	-266.5
LT	(9.9788)	(1.2492)		(60.166)	-268.0
GG	-1.2680	8068.25	.6844		-266.5
B2	1.0000	.7986	14731.3	.9742	-266.5
BR3	.9900	4.1936	2671.26	1.0000	-266.5
BR12	3.1856	6161.52	1.00	.2301	-267.1
LN	(9.9933)	(1.2704)			-268.0
W	.70115	42927.7	1.0000		-271.9
GA	1.0000	95736.5	.6182		-273.6
LOMAX	1.0000	38471.5	1.0000	1.6088	-269.2
EXP	1.0000	59183.3		1.0000	-275.7

NOTE: LL refers to the log of the likelihood function. Parameters in parentheses correspond to estimated parameters associated with the lognormal and log-t distributions.

Table 5
Estimated Quantiles
Unadjusted data

Model	α =.5	.1	.01
GB2(MLE)	16850	126600	1823000
LT(MLE)	21560	108830	426900
GG(MLE)	16850	126600	1823000
B2(MLE)	17610	120360	1345000
BR3(MLE)	17560	117790	1264000
BR12(MLE)	15610	142540	3297000
LN(MLE)	21880	111500	420300
W(MLE)	25450	141040	379000
GA(MLE)	31800	152890	350100
L(MLE)	20720	122490	635000
EXP(MLE)	41020	136270	272500
Normal Aprx			397200
Chebyshev			1373000
Normal Power			737500

Table 6

Quantiles

Model	α = .5	.10	.01
Population	115.7	196.7	330.5
GB2(MLE)	113.9	193.5	323.5
LT(MLE)	116.6	186.4	285.8
GG(MLE)	117.8	189.6	277.6
B2(MLE)	114.8	190.7	307.9
BR3(MLE)	113.3	193.4	364.7
BR12(MLE)	114.4	189.3	351.2
LN(MLE)	117.4	189.5	280.0
W(MLE)	122.4	201.9	269.5
GA(MLE)	120.4	190.8	264.4
EXP(MLE)	87.7	291.2	582.5
Normal Aprx	126.5	197.0	255.9
Chebyshev	203.3	298.3	669.9
Normal Power	144.0	207.4	330.0

Table 7
Cummins-Freifelder [1978] Severity Distribution

Grouped Format:	# of Losses
Less than 800	6
800 to 1442	15
1443 to 2093	10
2094 to 2820	7
2821 to 3696	9
3697 to 4845	6
4846 to 6527	7
6528 to 9471	3
9472 to 17124	6
17125 to 31158	6
31159 to 49803	0
49803 and over	5
	80

Table 8
Estimated Models
(Grouped Data)

Model	$a(\mu)$	$b(\sigma)$	p	q(d.f.)	LL	χ^2
GB2	1.5308	67.3722	71.1851	.5039	-17.6	2.6
LT	(7.9885)	(.9750)		(2.6064)	-21.3	10.0
GG	-1.5087	1095.8	.51210		-17.6	2.6
B2	1.0000	.00359	520181.	.9307	-17.8	3.0
BR3	.9687	.0247	57748.	1.0000	-17.9	3.1
BR12	2.9525	1095.4	1.0000	.2450	-17.6	2.8
LN	(8.1378)	(1.2378)			-21.6	10.3
W	.8275	5982.1	1.0000		-28.6	24.4
GA	1.0000	7896.6	.82068		-29.7	27.3
L	1.0000	5839.7	1.00000	1.5507	-24.4	15.4

NOTE: LL refers to the log of the likelihood function. Parameters in parentheses correspond to estimated parameters associated with the lognormal and log-t distributions.

Table 9
Severity Data
(Individual Observations)

Model	$a(\mu)$	$b(\sigma)$	p	q(d.f.)	LL	χ^2
GB2	3.9658	1097.4	.8524	.1866	-784.6	3.6
LT	(8.0159)	(1.0284)		(4.4832)	-791.6	10.6
GG	-1.0913	1656.2	.8401		-785.5	2.9
B2	1.0000	20.437	96.444	.9664	-785.6	3.1
BR3	.9918	31.991	62.194	1.0000	-785.6	3.2
BR12	3.5284	1062.0	1.0000	.2125	-784.6	3.4
LN	(8.2151)	(1.3490)			-794.7	11.4
W	.5810	7757.7	1.0000		-815.1	40.1
GA	1.0000	39513.	.4280		-830.9	68.1
L	1.0000	4640.1	1.0000	1.2941	-796.6	15.8
EXP	1.0000	16950.	1.0000		-859.0	110.0

NOTE: LL refers to the log of the likelihood function. Parameters in parentheses correspond to estimated lognormal and log-t parameters. The Chi-square tests are based on the groupings in Table 7, with the last three categories combined.

4. The Severity of Loss Distribution

The Cummins and Freifelder severity data consist of 80 fire claims.[9] These data are reported in grouped form in Table 7, and the ungrouped data are presented in Appendix B. Cummins and Freifelder found that the lognormal and gamma distributions did not have sufficiently heavy tails to describe the data. The log-t distribution was found to provide a better fit to the tail data.

Maximum likelihood estimation was used to estimate the members of the GB2 family. Both grouped and ungrouped data were used to analyze the impact of information lost through grouping. In estimating the models based upon the grouped data, the last three groups were combined as in Cummins and Freifelder. The results are reported in Tables 8 and 9. There appears to be very close agreement between the GB2, the two Burr distributions, the B2, and the GG for both the grouped and ungrouped data. The corresponding likelihood ratio tests are not statistically significant at the five percent level.

The Chi-square goodness of fit test does not provide the basis for rejecting either the LN or LT as being consistent with the data. However, the GB2, GG, B2, and the two Burr distributions provide statistically significant improvements over either the LN or the LT. The Chi-square test should be used with caution in this type of analysis, however, for two reasons: (1) It is not a very powerful test for goodness of fit of continuous distributions, and (2) it may fail to reject a distribution which provides an adequate fit in the body of the distribution but a very poor fit in the tail. Underestimating the tail can have serious consequences in an insurance context.

The Burr 12 results are very close to the GB2 results. Given the flexibility and relative simplicity of the Burr 12, it deserves to be given careful consideration in empirical work in this area. Another interesting observation is that while the estimated parameter values for the GB2 in Tables 8 and 9 appear quite different, the estimated distributions are very similar. It is not uncommon to find examples in which the likelihood surface is quite flat.[10]

5. Applications of Estimated Distributions.

The estimated frequency, severity, and total claims distributions are useful in numerous practical applications. The selection of distributional forms can have a significant impact on estimates of reinsurance premiums, tail fractiles (e.g., for MPY or ruin calculations), and other important statistics. In this section, we investigate the effects of model selection on the tails of the severity distribution, excess of loss reinsurance premiums, and simulated total claims distributions. We compare the GB2, Burr 12, and generalized gamma with the lognormal distribution. The first three distributions were selected because they fit the data best in terms of likelihood function values, while the lognormal was chosen because it has been used frequently in the prior literature. Recall that the lognormal could not be rejected using the chi-square test.

5.1. Tails of Severity Distributions. The tails (last 15 observations) of the estimated Burr 12 severity distributions based on grouped and ungrouped data are plotted in Figure 2. The tails of the two distributions differ noticeably, and the differences are large enough to have a significant impact on reinsurance premiums and other quantities. For the GB2 (not shown), the tails based on grouped and ungrouped parameter estimates are much more comparable than those of the Burr 12. Thus, one potential advantage of using the GB2 is that the results may be less sensitive to the type of data used in estimating the parameters.

The distribution function values, including the tails, of the GB2 and Burr 12 based on ungrouped data are virtually identical. For grouped data, the Burr 12 has a heavier tail than the GB2. In the latter case, the GB2 appears to fit the data slightly better than the Burr 12. The tails of the Burr 12, the generalized gamma, and the lognormal based on ungrouped data are presented in Figure 3. It is apparent that the tail of the lognormal is much too light to represent the data. Visual inspection also suggests that the Burr 12 fits the tail better than the generalized gamma.

5.2. Reinsurance Premiums. Reinsurance premium calculations were based on excess of loss reinsurance with both upper and lower limits. The calculations estimate the reinsurance company's expected severity. The expected severities would be multiplied by expected (total) frequency to give the reinsurance pure premium. The reinsurer's obligation for any given claim is the following:

$$(13) \qquad X_{RE} = \begin{cases} 0 & \text{for } X \le M \\ X - M & \text{for } M < X < U \\ U - M & \text{for } X \ge U \end{cases}$$

where X_{RE} = the reinsurer's obligation,

X = loss amount, and

M,U = lower and upper boundaries of the reinsurance layer.

The reinsurance severity expectations are given by the following formula:

$$(14) \qquad E(X_{RE}) = \int_M^U h(X)\ dF(X) + h(U)\ [1-F(U)] - h(M)\ [1-F(M)]$$

where $h(X)$ = a function of X: $h(X) = X$ for distributions based on X and $h(X) = \exp(X)$ for distributions based on $\ln(X)$.

For illustrative purposes, the upper boundary was set at \$1,000,000 and the lower boundary allowed to vary between $\ln M = 7$ (approximately \$1,000) and $1,000,000. The upper boundary point was chosen to be slightly higher than the largest observed claim (X_{MAX} = $626,000).

The differences between expected reinsurance severity based on grouped and ungrouped models tend to be quite noticeable. Consider, for example, the expected severities for the grouped and ungrouped the Burr 12 distributions, shown in Figure 4. The differences in the premiums range from about 15 percent for lower values of M to about 22 percent for higher values of M. Since the expected severities translate directly into pure premiums, the pure premium differences would be of the same magnitude. This is especially significant because the differences in the estimates could easily exceed the expected profit loading. Thus, the choice of the "wrong" distribution could result in inadvertently writing a policy at an expected loss (even ignoring investment income).

The importance of the choice of a severity model is reinforced by Figure 5, which presents the expected reinsurance severities based on ungrouped data for the GB2, Burr 12, generalized gamma, and lognormal distributions. The expected severities for the GB2 and Burr 12 are at least twice as large as those for the lognormal, over all values of M. The differences between the generalized gamma and the GB2 and Burr 12 are approximately 25 percent. This is especially significant in view of the fact that the generalized gamma distribution is not significantly different from the GB2 at the 10 percent level of significance (likelihood test). The difference is statistically significant at about the 17 percent level. Thus, in selecting a severity distribution model, it may be necessary to consider weaker significance criteria than the usual 5 or 10 percent levels.

5.3. Simulated Total Claims Distributions. The final application involved simulating the total claims distribution F(X) for various severity distributions. Two negative binomial frequency distributions were used:

Low Frequency: Failure parameter = 0.03714, Shape parameter = 95.1968

High Frequency: Failure parameter = 0.03714, Shape parameter = 3240.6435

The low frequency distribution is similar to the one used by Cummins and Freifelder [1978].[11] The expected number of claims per year for this distribution is 3.67, corresponding to the experience of the university supplying the fire insurance data. Since most insurance pools would have higher frequency, the high frequency distribution is also simulated. The expected number of claims in this case is 125 per year. This might correspond to a corporate self-insurance program. The number of frequency draws used in all of the simulations is 50,000.

The simulated total claims distributions are presented in Tables 10 through 12. Tables 10 and 11 give low frequency results for grouped and ungrouped data, while Table 12 gives the results for the high frequency distribution. The empirical moments of the estimated distributions also are given in the tables. These statistics should be interpreted carefully, since virtually none of the theoretical moments exist for the estimated generalized gamma, GB2, and Burr 12 distributions.

The results presented in Tables 10 through 12 reveal significant differences in the estimated quantiles depending upon the data used to estimate the parameters and the choice of model. For example, using ungrouped severity estimates and low frequency (Table 10), the 99-th percentiles of the total claims distributions range from $193,970 for the lognormal to $3,050,230 for the GB2. The GB2 and Burr 12 results are similar in Table 10 but differ substantially from the generalized gamma. Based on the grouped data (Table 11), the GB2 and generalized gamma results are comparable at the 99 percentile, but the Burr 12 yields a significantly larger value for X. The differences between the grouped and ungrouped low frequency estimates also are substantial. The GB2 performs best in terms of giving similar results using grouped and ungrouped data.

Due to the high costs of simulating the high frequency case, a full set of results was obtained only for the distributions based on ungrouped data. These results appear in Table 12. The overall conclusions are similar to those based on the low frequency case, i.e., there are substantial differences in the MPY values produced by the alternative models of F(x). For example, the 99th percentile of F(x) is $1,912,000 for the lognormal and $373,981,000 for the GB2. The generalized gamma also produces a 99 percent MPY much lower than that of the GB2. The Burr 12 results are comparable to those of the GB2. High-frequency simulation results with the Burr 12, generalized gamma, and lognormal reveal significant differences in MPYs based on the grouped and ungrouped data, similar to those observed in the low frequency case.

These results reinforce the conclusion of Cummins-Freifelder [1978] that the choice of a severity distribution can have a substantial impact on the estimation of MPY and ruin probabilities. The potential error in using the lognormal distribution is especially serious. Thus, the range of severity distributions considered in risk management applications should be significantly greater than in the past. The GB2 (or log-GB2) family of distributions should be sufficiently flexible for most insurance applications. The use of grouped rather than individual claim data also can lead to substantially different results. This reinforces the importance of collecting individual claim data, even when grouped estimates will be obtained as an alternative.

6. Summary

We have investigated several methods of estimating stochastic processes using the GB2 family of distributions. The GB2 has been shown to encompass many useful distributions and to provide a systematic way of considering them. The four parameters of this distribution provide sufficient flexibility to model most insurance loss and severity distributions. The GB2 provides at least as good a fit as any distribution which it subsumes, though it may be equaled when a special or limiting case adequately describes the process being modeled. In many cases, a three-parameter member of the family may be adequate. The Burr 12 is especially attractive in this regard because it has a closed form distribution function.

In modelling insurance loss distributions, we have shown that the choice of a distributional model for severity can have substantial effects on important statistics such as expected severities under reinsurance contracts and MPY estimates. The use of grouped individual claims data in estimating the severity distribution also can produce noticeably different results. In general, we believe that individual claim data are preferable to grouped data. However, especially in small samples, the use of grouped data provides a way to reduce the impact of potential outliers.

It is advisable to compute the relevant statistics using several different distributions in order to obtain an indication of the range of results that can be obtained using different models. When testing among alternative distributions, the chi-square test should be deemphasized and likelihood ratio tests should be used, with lower tolerances (e.g., 20 percent) than the usual 1, 5, or 10 percent because different tail behavior does not necessarily imply significant differences based upon comparisons involving the entire distribution.

Continuing improvement in computational hardware and software have made maximum likelihood techniques much more feasible than in the past. The gains from adopting these techniques and considering a broader range of probability distributions are potentially quite significant. The argument in favor of adopting these more general methodologies is especially strong when one considers the potential for error and the incomplete theoretical basis inherent in some techniques that have been used in the past to achieve computational simplicity.

APPENDIX A
The Pearson Family of Distributions and MPY

It will be convenient to define:

$$E(z) = \mu_z$$

$$\mu'_k(z) = E(z^k)$$

$$\mu_k(z) = E(z - \mu_z)^k,$$

for instances when these concepts exist.

Cummins and Wiltbank [1983], Kottas and Lau [1979] and Lau [1984] report some useful formulas which relate the first four moments of:

$$L = \sum_{i=1}^{n} x_i$$

about the sample mean to the similarly defined first four moments of "n" and the "x_i". These are given by:

(A.1) $$\mu_L = \mu_x \mu_n$$

(A.2) $$\mu_2(L)^2 = \mu_x \mu_2(n) + \mu_n \mu_2(X)$$

(A.3) $$\mu_3(L)^3 = \mu_x \mu_3(n) + \mu_n \mu_3(X) + 3\mu_x \mu_2(X) \mu_2(n)$$

(A.4) $$\mu_4(L) = \mu_x^4 \mu_4(n) + \mu_n \mu_4(X) + 4\mu_x \mu_3(X) \mu_2(n) + + 6\mu_x^2 \mu_2(X) [\mu_n \mu_2(n) + \mu_3(n)]$$

$$+ 3[\mu_2(X)]^2[\mu_n{}^2 - \mu_n + \mu_2(n)]$$

Equations (A.1)-(A.4) enable one to obtain the first four moments of "L" (and skewness and kurtosis) from the first four moments associated with frequency and severity. Computation of the these sample moments permits selection of the most appropriate Pearson distribution when utilized with these three steps:

A) Estimate the first four moments (ffm) of "x" and "n", using either maximum likelihood or sample moments.

B) Use equations (A.1-4) to calculate the ffm of "L".

C) Given the ffm of "L," use the Johnson, Nixon, Amos, Pearson [1963] tables or the equations of Bowman and Shenton [1979] to compute the required quantiles.

Common measures of skewness and kurtosis are given by β_1 and $\mu_3/(\mu_2)^{1.5}$ and $\beta_2 = \mu_4/\mu_2^2$.

There is no guarantee that the Pearson family will always provide a good fit to the L-distribution or that it will provide a better representation of "L" than other methods, such as those considered by Cummins and Freifelder [1978]. Wan and Lau [1981] noted that computation equations for the third and fourth moments are altered if the "x" are not independent. Such a lack of independence might occur in insurance computations, for example, if there was positive or negative contagion among data observations.

It should also be noted that for particular applications the data might be drawn from a population without the requisite number of theoretical moments to permit execution of method of moments calculations. Additionally, ready use of moments with Pearson tables requires that the moments be consistent with the range of values presented in the tables (a condition not always met, as is seen in the example used by Lau [1984]). The tables of Johnson, Nixon, Amos and Pearson (Johnson, et al. [1963]) require $0.0 \leq \beta_1 \leq 2.0$ and $0.0 < \beta_2 < 14.4$. The table values available for use with the equations of Bowman and Shenton (Bowman and Shenton [1979]) require $0.0 \leq \beta_1 \leq 2.0$ and $1.4 \leq \beta_2 \leq 15.8$.

With respect to this methodology one might note an early debate between Karl Pearson and R. A. Fisher. Their comments are summarized by Ord [1972, p.11] as resulting in the, "universally accepted" understanding that the method of moments is inefficient in the estimation of Pearson curves except when the curves are near the normal distribution. Moments were shown to be efficient only for distributions proportional to:

$$f(x) \propto \exp\ (a_1 x_1 + a_2\ x_2 + \ldots + a_n x_n).$$

Indeed, variance ratios suggest asymptotic efficiency (ratios in excess of 80 percent) only for $\beta_1 < 0.1$ and $2.62 \leq \beta_2 \leq 3.42$ when computing Pearsonian coefficients from the ffm.

Several methods of selecting among models have been considered. Two of these (the "k" and the "w" criteria) involve using sample moments to distinguish between members of the Pearson family. Ord [1972] compares these criteria and indicates that they are both inefficient when based on sample moments. The "w" criterion may have an advantage when the density starts at the origin.

Using the Kappa criterion (Elderton and Johnson [1969]) and the method of moments and the Pearson family (which includes both the B1 and B2) results in the selection of a B1:

$$B1(L;b,p,q) = \frac{(x/b)^{p-1}\ (1-x/b)^{q-1}}{b\ B(p,q)}$$
$$0 \leq x \leq b$$

with:

$b = 742995$, $p = .042568$ and $q = .70628$

for the unadjusted data.

This distribution (B1), however, has limited practical interest as a model for aggregate loss since it has an upper bound $\alpha = .01$.

TABLE 10

QUANTILES OF SIMULATED TOTAL CLAIM DISTRIBUTIONS
LOW FREQUENCY, 50,000 OBSERVATIONS
UNGROUPED SEVERITY*

QUANTILE	GB2	BURR12	GENERALIZED GAMMA	LOGNORMAL
0.50	17.68	17.71	18.50	21.57
0.60	25.24	25.15	25.31	28.21
0.70	37.66	37.34	35.92	37.12
0.80	63.67	62.69	55.36	50.16
0.90	155.75	150.61	109.36	75.35
0.91	175.79	170.49	121.10	79.35
0.92	203.25	196.08	137.39	84.20
0.93	241.46	232.45	156.11	89.93
0.94	294.48	283.01	181.39	96.56
0.95	375.00	358.52	221.18	104.32
0.96	497.30	473.29	279.12	115.52
0.97	729.09	689.19	370.84	129.39
0.98	1275.38	1196.90	560.30	152.38
0.99	3050.23	2826.37	1166.10	193.97
0.999	72841.54	64728.64	13458.12	435.48
MEAN	2535.61	2119.49	152.54	33.84
STD DEV	261895.11	214385.12	5070.00	44.78
SKEWNESS	166.68	165.65	118.76	7.83

TE: FREQUENCY IS NEGATIVE BINOMIAL WITH FAILURE PARAMETER 03714 AND SHAPE PARAMETER 95.1968. EXPECTED NUMBER OF IMS PER YEAR IS 3.67. TABLE ENTRIES ARE IN 000'S EXCEPT QUANTILES AND SKEWNESS.

VERITY DISTRIBUTION PARAMETERS ESTIMATED FROM UNGROUPED TA.

TABLE 11

QUANTILES OF SIMULATED TOTAL CLAIM DISTRIBUTIONS
LOW FREQUENCY, 50,000 OBSERVATIONS
GROUPED SEVERITY*

QUANTILE	GB2	BURR12	GENERALIZED GAMMA	LOGNORMAL
0.50	19.21	19.35	19.378	18.70
0.60	27.31	27.92	27.423	24.01
0.70	40.47	42.26	40.943	30.98
0.80	67.17	72.67	68.118	40.82
0.90	156.80	182.93	154.393	59.22
0.91	177.28	206.65	175.187	62.24
0.92	202.58	240.65	205.618	65.55
0.93	238.24	287.44	239.913	69.67
0.94	288.39	352.40	289.899	74.12
0.95	362.86	452.17	367.924	79.69
0.96	474.81	602.98	487.814	87.07
0.97	681.36	891.45	688.077	96.69
0.98	1168.71	1588.54	1142.17	111.02
0.99	2662.35	3882.27	2782.93	138.41
0.999	55309.24	99831.36	49193.647	284.25
MEAN	1578.79	3883.75	650.778	27.11
STD DEV	151354.39	415080.41	38923.73	30.91
SKEWNESS	163.15	168.45	136.301	5.38

NOTE: FREQUENCY IS NEGATIVE BINOMIAL WITH FAILURE PARAMETER 0.03714 AND SHAPE PARAMETER 95.1968. EXPECTED NUMBER OF CLAIMS PER YEAR IS 3.67. TABLE ENTRIES ARE IN 000'S EXCEPT FOR QUANTILES AND SKEWNESS.

*SEVERITY DISTRIBUTION PARAMETERS ESTIMATED FROM GROUPED DATA.

TABLE 12

QUANTILES OF SIMULATED TOTAL CLAIM DISTRIBUTIONS
HIGH FREQUENCY, 50,000 OBSERVATIONS
UNGROUPED SEVERITY*

QUANTILE	GB2	BURR12	GENERALIZED GAMMA	LOGNORMAL
0.50	3023.76	2893.43	1857.77	1116.58
0.60	3903.74	3711.30	2171.01	1175.80
0.70	5409.47	5099.43	2662.45	1244.25
0.80	8469.46	7893.35	3574.66	1332.23
0.90	18614.66	17042.23	6129.56	1467.12
0.91	21114.32	19268.32	6718.28	1486.49
0.92	24782.75	22555.95	7408.53	1509.26
0.93	29468.11	26717.69	8232.88	1534.46
0.94	35786.90	32330.38	9336.65	1563.81
0.95	44956.99	40467.63	10875.72	1598.11
0.96	58537.12	52494.64	13301.50	1642.23
0.97	86803.76	77472.09	17613.00	1694.32
0.98	147382.83	130296.87	26626.59	1775.58
0.99	373980.84	325678.38	54672.50	1911.62
0.999	7836815.00	6552029.50	628293.38	2488.61
MEAN	67399.30	180539.81	7878.91	1148.67
STD DEV	2238155.45	21393692.13	238635.71	256.04
SKEWNESS	70.67	200.71	133.11	1.46

NOTE: FREQUENCY IS NEGATIVE BINOMIAL WITH FAILURE PARAMETER 0.03714 AND SHAPE PARAMETER 3240.6435. EXPECTED NUMBER OF CLAIMS PER YEAR IS 125. TABLES ENTRIES ARE IN 000'S EXCEPT FOR QUANTILES AND SKEWNESS.

*SEVERITY DISTRIBUTION PARAMETERS ESTIMATED FROM UNGROUPED DATA.

FIGURE 2
BURR12 TAILS: GROUPED & UNGROUPED

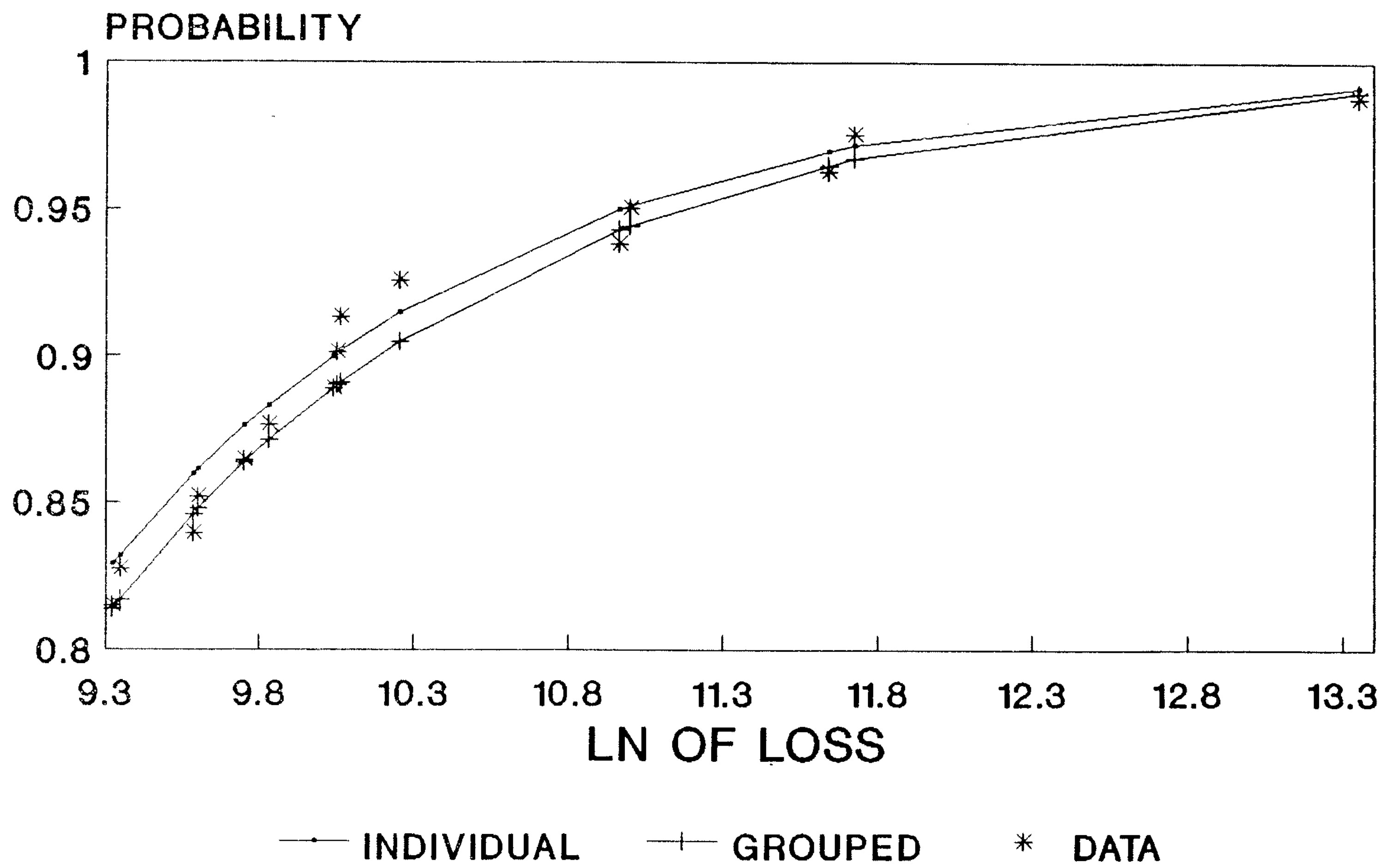
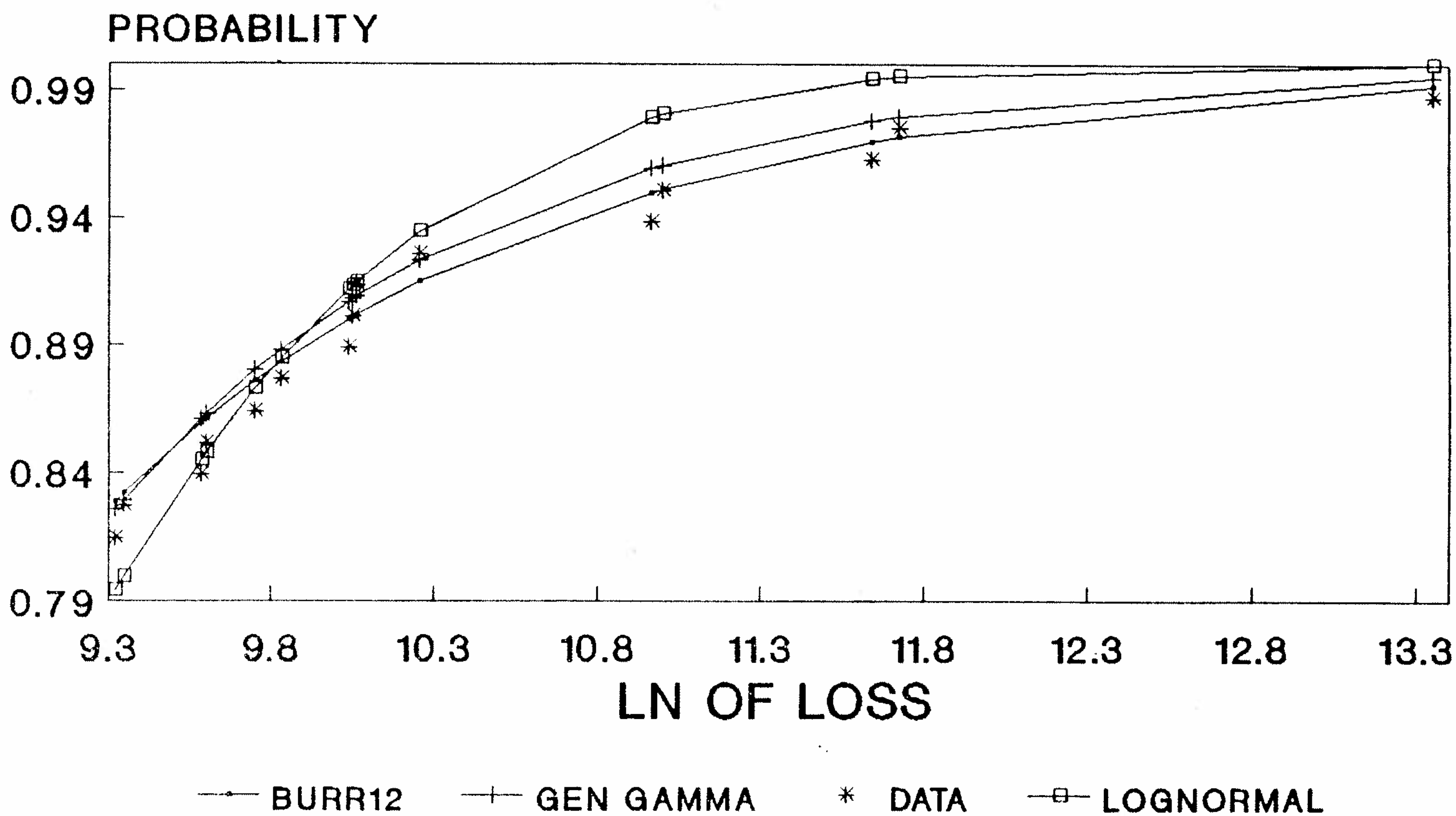
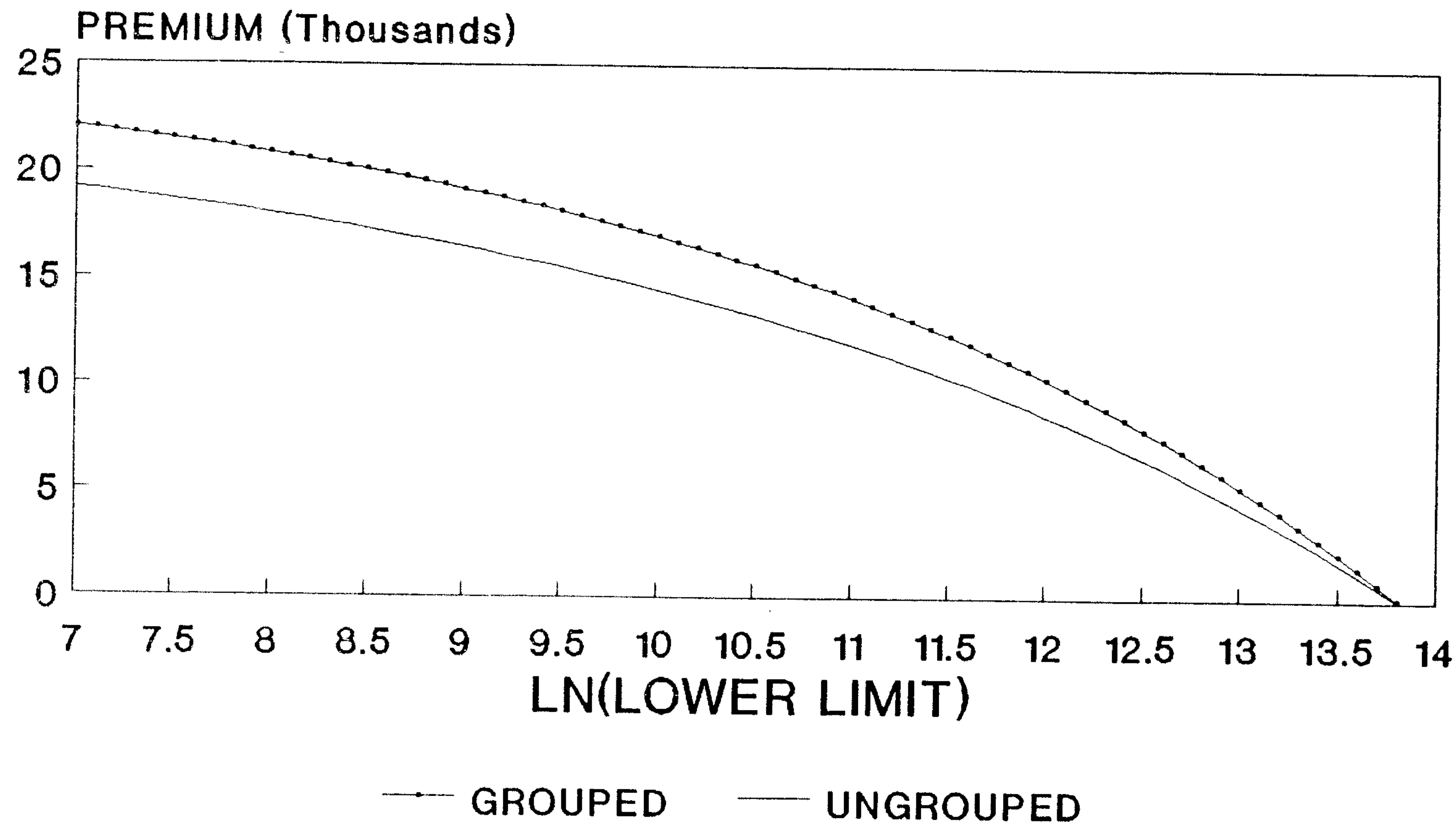


FIGURE 3
BURR 12, GEN GAMMA, & LOGNORMAL TAILS



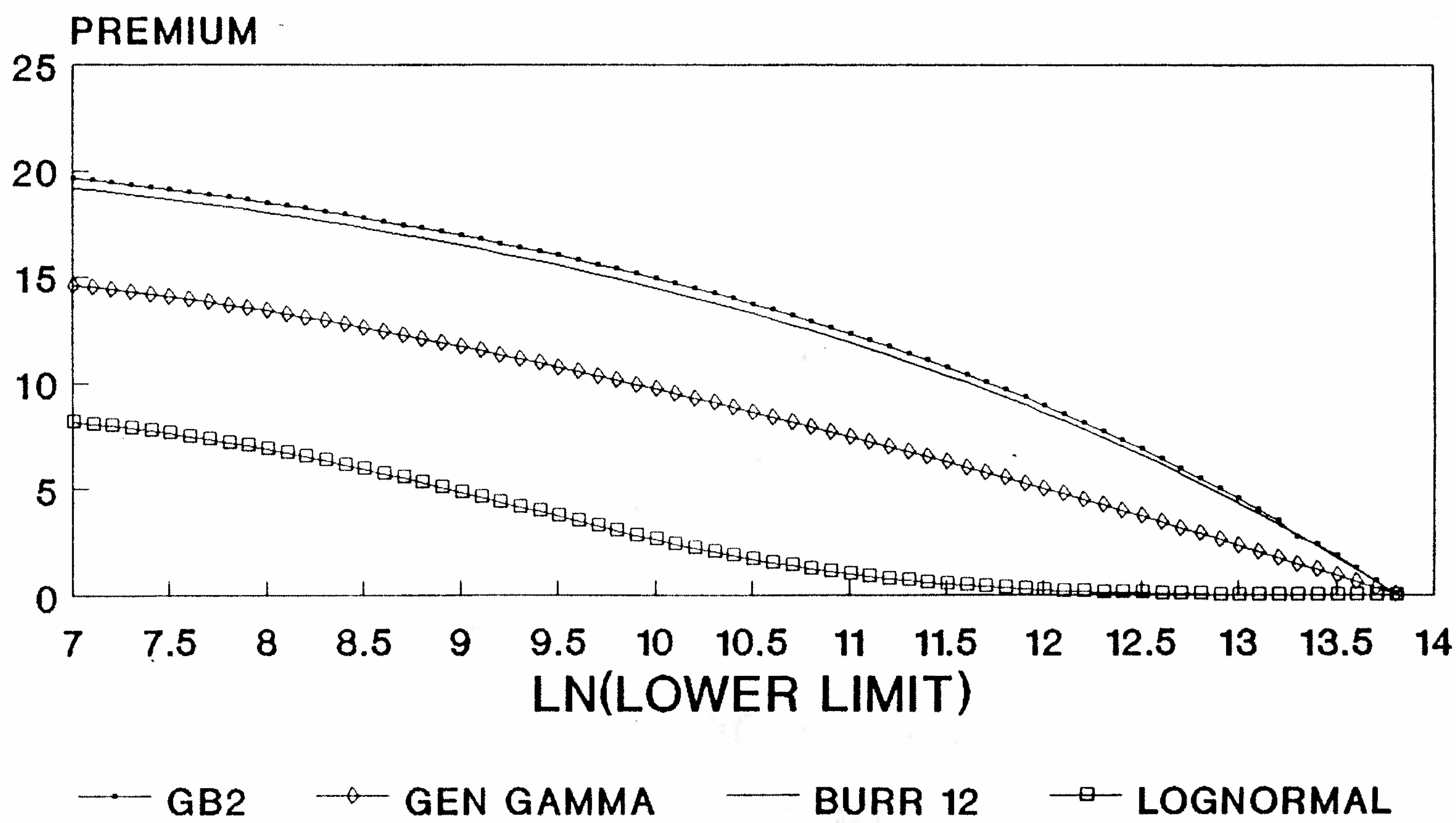
UNGROUPED DATA

FIGURE 4
BURR 12
EXPECTED REINSURANCE SEVERITY



UPPER LIMIT = $1 MILLION

FIGURE 5
GB2, BURR 12, GEN GAMMA, & LOGNORMAL EXPECTED REINSURANCE SEVERITY



UPPER LIMIT = $1,000,000

APPENDIX B

Cummins-Freifelder [1978] Severity Data: Individual Observations

290.4	1248.49	2202.96	3941.3	10560.1
537.19	1268.24	2222.8	4017.01	11179.54
756.8	1284.56	2255.72	4100	11461.39
769.19	1363.85	2274.61	4166.98	14538.13
787.69	1436.2	2328.64	4355.02	14789.81
796.18	1445.96	2384.37	5117.93	17186.09
933.62	1469.48	2847.83	5335.96	18582.57
967.97	1507.47	2947.04	5453.02	22857.33
1010.56	1662.36	2948.35	5568.96	23177.85
1017.4	1674.58	3036.51	5761.83	23446.13
1033.49	1690.91	3287.68	6161.81	28409.82
1034.33	1739.96	3331.62	6348.69	57612.82
1056.93	1776.56	3416.67	6859.37	59582.78
1124.09	1932.09	3604.66	7972.2	113164.7
1165.73	1975.89	3671.16	8028.32	123228.9
1217.64	2099.79	3739.3	10047.22	626402.8

REFERENCES

Abramowitz, Milton and Irene A. Stegun. [1965]. Handbook of Mathematical Functions. New York: Dover Publications, 1965.

Aiuppa, Thomas A. [1986]. "Maximum Probable Yearly Aggregate Loss: A Pearson Approximation. Forthcoming in Journal of Risk and Insurance.

Beard, R.E., T. Pentikainen and E. Pesonen [1965]. Risk Theory, Methuen, London.

Benckert, L.G. [1962]. "The Log-Normal Model for the Distribution of One Claim," ASTIN Bulletin, Vol. 2, Part I, pp. 9-23.

Bowman, K.O. and L.R. Shenton [1979]. "Approximate Percentage Points for Pearson Distributions," Biometrika, Vol. 66(1), pp. 147-151.

________ [1979]. "Further Approximate Pearson Percentage Points and Cornish-Fisher," Communications in Statistics, Vol. B8(2), pp. 231-244.

Buhlmann, Hans [1970]. Mathematical Methods in Risk Theory. New York: Springer-Verlag.

Cummins, J.D. and L.R. Freifelder [1978]. "A Comparative Analysis of Alternative Maximum Probable Yearly Aggregate Loss Estimates," Journal of Risk and Insurance, Vol. 45(1), pp. 27-52.

________, and Laurel J. Wiltbank [1983]. "Estimating the Total Claims Distribution Using Multivariate Frequency and Severity Distributions," Journal of Risk and Insurance, Vol. 50, pp. 377-403.

________ [1984]. "A Multivariate Model of the Total Claims Process," ASTIN Bulletin, Vol. 14, pp. 45-52.

Edlefson, Lee E. and Samuel D. Jones [1986]. Gauss. Kent, WA: Applied Technical Systems.

Elderton, W.P. [1938]. Frequency Curves and Correlation. Cambridge University Press.

________ and N.L. Johnson [1969]. Systems of Frequency Curves. London: Cambridge University Press (1.3,1.7,1.11).

Ferreira, Joseph, Jr. [1970]. Quantitative Models for Automobile Accidents in Insurance. U.S. Department of Transportation Automobile Insurance and Compensation Study. Washington, D.C.: U.S. Government Printing Office.

Hall, W.J. and J.A. Willner [1981]. "Mean Residual Life." In M. Csorgo, et al., eds., Statistics and Related Topics. Amsterdam: North-Holland Publishing Co.

Head, George L. [1971]. Insurance to Value. Homewood, Ill.: Richard D. Irwin.

Heckman, P.E. and G.G. Meyers [1983]. "The Calculation of Aggregate Loss Distributions From Claim Severity and Claim Count Distributions," Proceedings of the Casualty Actuarial Society, Vol. 70.

Hewitt, C.C. [1970]. "Credibility for Severity," Proceedings of the Casualty Actuarial Society, Vol. 57.

Hogg, R.V. and S.A. Klugman [1984]. Loss Distributions. New York: John Wiley and Sons.

IMSL [1985]. International Mathematics and Statistics Library. Houston, TX.

Johnson, Norman L. and Samuel Kotz [1970]. Continuous Univariate Distributions--1. New York: John Wiley & Sons.

Johnson, N.L., E. Nixon, D.E. Amos and E.S. Pearson [1963]. "Table of Percentage Points of Pearson Curves," Biometrika, Vol. 50, pp. 459-498.

Kendall, M.G. and A. Stuart [1969]. The Advanced Theory of Statistics. Vol. I. London: Charles Griffin.

Kottas, J.F. and H.S. Lau [1979]. "A Realistic Approach for Modeling Stochastic Lead Time Distributions," AIIE Transactions, Vol. 11(1), pp. 54-60.

Lau, Hon-Shiang [1984]. "An Effective Approach for Estimating the Aggregate Loss of an Insurance Portfolio," The Journal of Risk and Insurance, Vol. 50, pp. 21-30.

McDonald, James B. [1984]. "Some Generalized Functions for the Size Distribution of Income," Econometrica, Vol. 52, pp. 647-663.

________, and Dale O. Richards [1987]. "Model Selection: Some Generalized Distributions," Communications in Statistics, Vol. 16, pp. 1049-1074.

________, and Richard J. Butler [1987]. "Some Generalized Mixture Distributions with an Application to Unemployment Duration," The Review of Economics and Statistics, Vol. 69, pp. 232-240.

Mandelbrot, B. [1964]. "Random Walks, Fire Damage and Other Paretian Risk Phenomena, "Operations Research, Vol. 12.

Ord, J.K. [1972]. Families of Frequency Distributions. New York: Hafner Publishing Company.

Panjer, H.H. [1981]. "Recursive Evaluation of a Family of Compound - Distributions," ASTIN Bulletin, Vol. 12, pp. 22-30.

Paulson, A. S. [1984]. "Estimating the Total Claims Distribution In Property-Liability Insurance," working paper, S. S. Huebner Foundation, University of Pennsylvania.

________ and R.A. Dixit [1986]. "Cash Flow Simulation Models for a Large Property-Liability Insurance Pool," paper presented at First International Conference on Insurance Solvency, Philadelphia, Center for Research on Risk and Insurance, University of Pennsylvania.

________ and N.J. Faris [1985]. "A Practical Approach to Measuring the Distribution of Total Annual Claims." In J. D. Cummins, ed., Strategic Planning and Modelling in Property-Liability Insurance. Norwell, MA: Kluwer Academic Publishers.

Pearson, E.S. and H.O. Hartley [1972]. Biometrika Tables for Statisticians. Vol.II. Cambridge: Cambridge University Press.

Roy, Yves and J. David Cummins [1985]. "A Stochastic Simulation Model for Reinsurance Decision Making By Ceding Companies." In J. D. Cummins, ed., Strategic Planning and Modelling in Property-Liability Insurance. Norwell, MA: Kluwer Academic Publishers.

Seal, Hilary L. [1969]. Stochastic Theory of a Risk Business. New York: John Wiley and Sons.

Shpilberg, D.C. [1977]. "The Probability Distribution of Fire Loss Amount," Journal of Risk and Insurance, Vol. 44, pp. 103-115.

Tadikamalla, Pandu R. and John S. Ramberg [1975], "An Approximate Method for Generating Gamma and Other Variates," Journal of Statistical Computation and Simulation, Vol. 3, pp. 275-282.

Van der Laan, B.S. [1988]. Modelling Total Costs of Claims on Non-Life Insurances. Delft, Holland: Eburon Publisher.

Venter, Gary C. [1984]. "Transformed Beta and Gamma Functions and Aggregate Losses," reprinted from Vol. 71, Proceedings of the Casualty Actuarial Society. Boston: Recording and Statistical Corporation, 1984.

Wan, W.X. and H.S. Lau [1981]. "Formulas for Computing the Moments of Stochastic Lead Time Demand," AIIE Transactions, Vol. 13, pp. 281-282.

FOOTNOTES

1. This includes the Pearson Type IV, which statisticians have traditionally considered intractable (see, for example, Johnson and Kotz [1970, p. 12]).

2. The Lomax is seen to be a mixture of an exponential where the scale parameter is distributed as an inverse gamma. The Lomax has been referred to as a compound exponential distribution. The beta of the second kind (B2) is a mixture and is also known as a generalized Pareto.

3. A cross section of risk pools might also be used, but this situation is less likely to arise in practice.

4. If the aggregate approach is adopted, great care must be taken to allow for trends, cycles, and other sources of non-stationarity that may be present in the data. Of course, the non-stationarity problems will be reduced if data are available on a quarterly or monthly basis, provided the pool size is sufficient. Methods to adjusting for trends, inflation, and other data problems are discussed in Cummins and Freifelder [1978].

5. The equations that must be solved to yield method of moments estimators for the GB2 are the following (moment formulas for other members of the family are given in McDonald and Richards [1987]):

$$E_{GB2}(X^h) = \frac{b^h \, B(p + h/a, q-h/a)}{B(p,q)},$$

$$= \sum_{t=1}^{n} X_t^h/n$$

$$= \mu_h'(X), \; h = 1,2,3,4,$$

for a, b, p and q, where $\mu'(X)$ denotes the estimated h^{th} order moment of X about the origin. Cummins and Wiltbank [1983], Kottas and Lau [1979] and Lau [1984] report some formulas which relate the first four moments of aggregate loss to the first four moments of severity and frequency.

6. Empirical moments also have provided the basis for selecting members of the Pearson family. Applications of the Pearson family in modelling aggregate losses are discussed by Lau [1984] and Aiuppa [1986]. A brief summary of this approach and related issues are presented in Appendix A.

7. The table corresponds to the unadjusted data (unadjusted for the varying number of exposure units or other factors discussed by Cummins and Freifelder [1978] and reported in table 3). Thus the results in this section abstract from adjustment problems and are only intended to

illustrate estimation of the MPY.

8. The estimates provided by this method tend to deteriorate for higher skewnesses (Beard, Pentikainen, and Pesonen [1984]).

9. The data cover several years and have been adjusted to a common time point using a claims cost index maintained by the university from which the claims were obtained. The data are described in more detail in Cummins and Freifelder [1978].

10. Depending upon the starting values chosen in the maximum likelihood estimation, it is possible to "converge" to different parameter estimates for the GB2 from the same data set. For example, the authors obtained the following alternative set of GB2 parameters from the ungrouped data:

    $a = 1.1132 \quad b = 20.9366 \quad p = 126.9381 \quad q = 0.8202$

    This parameter set yielded a log-likelihood function value of -785.5, which is not as good as the results reported in Table 9, but the estimated distribution functions are virtually identical.

11. These parameters are comparable to those in the Cummins-Freifelder article [1978, Table 2]. The shape parameter for one exposure unit is identical to the one used by Cummins and Freifelder (0.2978). The failure parameter used here (0.03714) is slightly larger than the one used by Cummins and Freifelder (0.0371). The latter parameter does not appear in Table 2 of Cummins-Freifelder because they used a different but mathematically equivalent formulation of the negative binomial. For purposes of comparison, the failure parameter (0.0371) has been computed from the second parameter given in Table 2 of Cummins-Freifelder. The shape parameter used in the low frequency simulations in the present article represents approximately 320 exposure units while the high frequency parameter implies about 10,882 exposure units.