THE EFFECTS OF PARAMETER UNCERTAINTY IN THE EXTREME EVENT FREQUENCY-SEVERITY MODEL

JAKUB M. BOROWICZ [†]

and

JAMES P. NORMAN[‡]

ABSTRACT

In this paper we analyse parameter uncertainty in the extreme event frequency-severity model. Three methods are compared: classical asymptotic statistics, bootstrapping and a Bayesian approach. The Bayesian method makes use of Markov Chain Monte Carlo techniques, which have recently become more feasible through the use of specialised software and fast computers. These approaches are demonstrated using a data set of large Danish fire losses, previously analysed by several authors. The effects of parameter uncertainty on capital requirements and the price of an Excess of Loss reinsurance contract are investigated in a Dynamic Financial Analysis (DFA) framework. Our results show that, in this case, parameter uncertainty has a significant effect and should not be ignored in DFA models.

Keywords: Bayesian, Parameter Uncertainty, Large Claims, Extreme Values, DFA, Generalised Pareto Distribution.

[†] Jakub.Borowicz@emb.co.uk

[‡] James.Norman@emb.co.uk

Address: EMB Consultancy LLP, Link House, Cambridge, UK, CB2 5LT

Summary

1. Introduction

Parameter uncertainty arises from the difficulty in assessing the true parameter values of a distribution from which a finite sample has been taken. Parametric distributions are often fitted to historical claims severity and frequency observations for use as a guide for future claims. Traditionally, parameters have been estimated as deterministic values using, for example, the method of moments or maximum likelihood estimation (MLE[1]). However, it should be remembered that these estimators are themselves realisations of random variables – if two finite samples are taken from an identical distribution, the maximum likelihood estimates of the parameters for both samples are likely to be different, and in general neither will give the true parameter value. Using fixed parameters in the frequency-severity model for claims will lead to misestimation of risk, as the estimated parameters may be based on little experience.

An analysis of the uncertainty in parameters has been mentioned recently by regulatory bodies as a factor to consider in assessing capital. Common methods of assessing parameter uncertainty include "classical" statistical methods such as the asymptotic normality of maximum likelihood estimators, and empirical approaches such as non-parametric bootstrapping.

Parameter uncertainty in various aspects of insurance has been discussed previously by many authors (see, for example [14],[3],[16]). A more thorough study of parameter uncertainty has been made possible with the increase in power of computers and the availability of professional software. Markov Chain Monte Carlo methods allow simulation directly from the posterior distribution of parameters.

In our study we use the set of Danish fire losses as analysed in [19],[15],[18]. We have restricted our attention to the generalised Pareto distribution for modelling the severity and Poisson for the frequency. We find that it is necessary to upper-truncate the predictive distribution to avoid simulating unrealistically large losses. Parameter uncertainty is shown to increase the capital requirement and estimated reinsurance premiums.

[1] Throughout this paper, the abbreviation "MLE" is used interchangeably for Maximum Likelihood Estimation, Maximum Likelihood Estimate and Maximum Likelihood Estimator.

2. Model

2.1 Frequency - Severity

We model large claims using a frequency-severity approach, so that the annual aggregate loss is a sum of a random number of individual losses:

$$S = Y_1 + \ldots + Y_N \,. \tag{2.1.1}$$

We assume that claim amounts are a sequence of iid. random variables, and that N is also a random variable independent of this sequence. Common choices for the frequency-severity distributions are the Poisson distribution for the claim numbers and the generalised Pareto distribution (GPD) for the claim amounts. We will expand on the analysis of the Poisson-GPD model by incorporating parameter uncertainty. Previous analysis of parameter uncertainty in this model can be found in [14] where the bootstrap technique is used. Here, we will use a Bayesian approach. The method we demonstrate is generic, and can be applied to produce distributions of parameters in most situations where point parameters are estimated by maximising a likelihood function.

2.2 Severity

The generalised Pareto distribution has the following CDF [7]:

$$F(x) = \begin{cases} 1 - \left(1 + \dfrac{\xi(x-\tau)}{\sigma}\right)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0, \\ 1 - \exp\left(-\dfrac{x-\tau}{\sigma}\right) & \text{if } \xi = 0, \end{cases} \tag{2.2.1}$$

where ξ is a shape parameter, σ is the scale parameter and τ a lower threshold.

The use of the generalised Pareto distribution in modelling the severity is motivated by classical extreme value theory [7], which states that, for a large class of distributions, the distribution of exceedances over a high threshold becomes approximately GPD. A possibly undesirable feature of this distribution is that it can give unreasonably high losses with finite probability, a property that is exacerbated when considering parameter uncertainty. For this reason, we instead model severity using a generalised Pareto distribution that has been truncated from above at some large value of loss T. The truncation point could represent, for example, the maximum policy limit of the risks in the portfolio.

The cumulative distribution function for such a truncated distribution is :

$$G(x) = \frac{F(x)}{F(T)}, \quad \tau < x < T \ . \tag{2.2.2}$$

Our model is calibrated to historical data. The data we use is the set of Danish fire losses from [19], which has been analysed by a number of other authors. We have not performed any adjustments to the data. The upper threshold *T*, in the absence of any policy limit data, has been chosen to be four times the maximum observed loss in the data set. The parameters ξ (shape) and σ (scale) were estimated using the maximum likelihood method, for which numerical routines are needed.

We also apply goodness-of-fit tests to verify the model assumption. The χ^2-goodness-of-fit test is a commonly used test to use when assessing goodness-of-fit. However, this test cannot be applied for small data sets [4]. Two alternative tests that do not suffer from this drawback are the Kolmogorov-Smirnov test and the Anderson-Darling test. However, the distribution of the statistic for these tests is not known in the case when parameters have been estimated from the data [4]. Here, we use Monte Carlo simulation to obtain the approximate distribution of the test statistic for each selected threshold τ.

Our results for the fitted maximum likelihood estimates of the parameters, together with the results of the goodness-of-fit tests, are shown in Table 1. The values of the lower threshold are explicitly chosen to maintain consistency with earlier authors.

Table 1: Fitted maximum likelihood parameters for the generalised Pareto distribution at selected thresholds in millions of Danish Krone (DKM), with goodness-of-fit results.

Threshold (DKM)	σ	ξ	χ^2 p-value	KS p-value	AD p-value
5	3.79	0.64	0.03	0.02	0.02
10	6.96	0.50	0.43	0.88	0.72
20	9.49	0.73	-	0.63	0.90

2.3 Frequency

Maximum likelihood estimation for the Poisson distribution is a simple exercise, as the MLE of the Poisson parameter is the sample mean, $\bar{x}$. We fit the Poisson distribution to the number of losses above the previously selected threshold occurring in each of the 11 accident years. The Kolmogorov-Smirnov and Anderson-Darling tests mentioned in section 2.2 cannot be applied to discrete distributions such as the Poisson. However, the χ^2 goodness-of-fit test may still be applied. The results are shown in Table 2.

We have also performed a test for under/over dispersion, since the Poisson distribution has variance equal to the mean. The Böhning test [1] is one such test. The hypothesis of unit dispersion is not rejected.

Table 2: Fitted maximum likelihood parameters for the Poisson distribution at selected thresholds, with goodness-of-fit results.

Threshold (DKM)	Sample mean	Sample variance	χ2 p-value	Böhning p-value
5	23.09	42.49	0.62	0.06
10	9.91	8.29	0.37	0.71
20	3.27	5.62	0.01	0.11

3. Parameter Uncertainty

3.1 Introduction and Methods

The frequency-severity model above has been calibrated to limited historical data; the true parameter values are uncertain. The theory of maximum likelihood estimation states that, for large sample sizes, n, maximum likelihood estimators, $\hat{\Theta} = \left(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k\right)$, are approximately distributed as a multivariate normal [11],

$$\hat{\Theta} n^{1/2} \xrightarrow{d} N\left(\hat{\Theta} n^{1/2}, \left[\mathbf{I}_1(\Theta)\right]^{-1}\right), \quad \text{as} \quad n \to \infty, \qquad (3.1.1)$$

with covariance matrix given by the inverse of the Fisher information matrix.

The Fisher information matrix, $\mathbf{I}_1(\Theta)$, is defined for a k-dimensional parameter vector as:

$$\mathbf{I}_1(\Theta) = -\frac{1}{n}E\begin{bmatrix} \frac{\partial^2}{\partial\theta_1^2}\ln L & \frac{\partial^2}{\partial\theta_1\partial\theta_2}\ln L & \cdots & \frac{\partial^2}{\partial\theta_1\partial\theta_k}\ln L \\ \frac{\partial^2}{\partial\theta_2\partial\theta_1}\ln L & \frac{\partial^2}{\partial\theta_2^2}\ln L & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial\theta_k\partial\theta_1}\ln L & \cdots & \cdots & \frac{\partial^2}{\partial\theta_k^2}\ln L \end{bmatrix}, \tag{3.1.2}$$

where L is the likelihood function.

For the generalised Pareto distribution, one can show that covariance matrix of the MLE parameters is[2],

$$\frac{1}{n}\left[\mathbf{I}_1(\xi,\sigma)\right]^{-1} = \frac{1}{n}\begin{pmatrix} (1+\xi)^2 & -\sigma(1+\xi) \\ -\sigma(1+\xi) & 2\sigma^2(1+\xi) \end{pmatrix}. \tag{3.1.3}$$

This expression uses the true values of the parameters. However, it is common to estimate the covariance matrix using the maximum likelihood estimates [11].

For many distributions, the Fisher information matrix is difficult to calculate, and so approximations must be used. For the upper-truncated generalised Pareto distribution, we approximate the variance-covariance matrix using numerical differentiation [11]:

$$\frac{1}{n}\left[\mathbf{I}_1(\xi,\sigma)\right]^{-1} \approx -\begin{bmatrix} \frac{\partial^2}{\partial\xi^2}\ln L(\xi,\sigma) & \frac{\partial^2}{\partial\xi\,\partial\sigma}\ln L(\xi,\sigma) \\ \frac{\partial^2}{\partial\xi\,\partial\sigma}\ln L(\xi,\sigma) & \frac{\partial^2}{\partial\sigma^2}\ln L(\xi,\sigma) \end{bmatrix}^{-1}. \tag{3.1.4}$$

The asymptotic normality assumption can provide a useful guide to the variance of the fitted parameters. However, under the normality assumption the parameters can become negative, which is invalid for a generalised Pareto distribution.

[2] We note here that the off-diagonal elements of the covariance matrix presented here differ in sign from that presented in [15].

Another approach to assessing parameter uncertainty is bootstrapping [4]. In non-parametric bootstrapping, the sampling distribution of an estimator is approximated by resampling with replacement from the original sample. The idea behind non-parametric bootstrapping is that since we lack any information about a population from which the sample has been taken, the values in that sample are the best guide to the true distribution. This method was applied to the frequency-severity model fitted to the Danish fire data in [14]. The disadvantage of this method is that parameter values are restricted within a range, and multi-modal distributions for parameters can be produced. Further, bootstrapping can be computationally intensive. An alternative is to use a Bayesian approach, which avoids some of the drawbacks of bootstrapping or the normal MLE approximation.

A critical feature of Bayesian analysis is the choice of prior distribution for the parameters. In our analysis we choose uniform priors [13], as we do not assume any prior knowledge of parameter values, and we allow all the information about parameters to come from the data. We give a brief summary of Bayesian analysis in section 3.2. For a more formal introduction, the reader is directed to [2].

3.2 Bayesian Analysis

Suppose we take observations, $\underline{y}$, from a random variable Y and wish to make inferences about another random vector, $\Theta = (\theta_1, \theta_2, ..., \theta_k)$, where Θ is drawn from a distribution with density function $p(\Theta)$. It can be seen that:

$$p(\Theta \mid \underline{y}) = \frac{p(\underline{y} \mid \Theta)p(\Theta)}{p(\underline{y})} = \frac{p(\underline{y} \mid \Theta)p(\Theta)}{\int p(\underline{y}, \Theta)d\Theta}, \tag{3.2.1}$$

where $p(\Theta)$ is the assumed prior distribution of the unknown parameters, while $p(\Theta \mid \underline{y})$ is the posterior distribution given the prior $p(\Theta)$ and the data $\underline{y}$.

The maximum likelihood estimation method and Bayesian analysis are closely related. Let $L(\Theta \mid \underline{y})$ be the assumed likelihood function. Under MLE, we have computed the maximum value of the likelihood function, L, as a function of Θ given the data, $\underline{y}$. A useful feature of MLE is its large-sample properties, namely that when the sample size is sufficiently large we can assume approximate multivariate normality of the maximum likelihood estimators. These convenient attributes do not necessarily hold for small samples, such as those analysed here.

An alternative way to proceed is to start with some initial knowledge about the distribution of the unknown parameters $p(\Theta)$. From Bayes' Theorem, given the likelihood and the prior distribution, we produce a posterior distribution:

$$p(\Theta \mid \underline{y}) = \frac{1}{p(\underline{y})} p(\underline{y} \mid \Theta) p(\Theta) = \text{constant} \times \text{likelihood} \times \text{prior}\,. \tag{3.2.2}$$

Furthermore, as $p(\underline{y} \mid \Theta) = L(\Theta \mid \underline{y})$ is just the likelihood function and $1/\mathrm{p}(\underline{y})$ is a constant with respect to Θ, we can write the posterior distribution as:

$$p(\Theta \mid \underline{y}) \propto L(\Theta \mid \underline{y}) p\,(\Theta), \tag{3.2.3}$$

where the constant of proportionality normalises $p(\Theta \mid \underline{y})$ to 1, and can be obtained by integration.

Under uniform prior distributions[3] for Θ, the posterior distribution is simply proportional to the likelihood:

$$p\left(\Theta \mid \underline{y}\right) \propto L\left(\Theta \mid \underline{y}\right). \tag{3.2.4}$$

A criticism of Bayesian analysis is that the posterior distribution is dependent on the choice of prior. We do not investigate the influence of different choices of prior distribution here. However, if the data contains enough information, we expect that the posterior distribution will not be greatly effected by any reasonable choice of prior.

Another important concept of Bayesian analysis is that of the predictive distribution. This is the conditional distribution of a new observation given the historical data [11]. The predictive distribution can be calculated as the mixture of the posterior distribution and the model distribution. It is the predictive distribution which should be used when forecasting future claims experience.

[3] We note here that if the distribution is defined as uniform on the unbounded range, it cannot integrate to unity. Strictly speaking, in our study we use locally uniform priors [13] that change little over the region in which the likelihood is of interest and do not take very large values outside that region.

3.3 Parameter Uncertainty for the Frequency Distribution

Consider n observations, $\underline{y}$, of the random variable Y, where Y represents the number of claims per year. For the Poisson distribution, we are fortunate that the posterior distribution of the parameter μ is exactly solvable. In fact, with uniform prior, it can be shown that[4]

$$\mu \mid \underline{y} \sim \mathrm{Gamma}(\alpha, \beta), \tag{3.3.1}$$

where $\alpha = \sum_{i=1}^{n} y_i + 1$ and $\beta = 1/n$.

For other choices of frequency distribution, it may not be possible to write the posterior distribution in such a simple form, in which case we would have to use the Gibbs/ARMS algorithm [10],[9], as will be necessary for the severity distribution. For the Danish fire data, the parameters of the gamma distribution are given in Table 3.

Table 3: Parameters of the posterior Gamma distribution with its mean, standard deviation (SD) and Coefficient of Variation (CoV).

Threshold (DKM)	α	β	Mean	SD	CoV
5	255	0.09	23.18	1.45	0.06
10	110	0.09	10.00	0.95	0.10
20	37	0.09	3.36	0.55	0.16

From here on, we restrict attention to the case where the threshold is 10 DKM.

3.4 Parameter Uncertainty for the Severity Distribution

The posterior distribution of the parameters of the truncated generalised Pareto distribution is not recognisable as any standard distribution. If we consider m observations, $\underline{x}$, of the random variable X, the log-likelihood is:

[4]The convention used is such that the expected value of Gamma(α,β) is $\alpha\beta$ and the variance is $\alpha\beta^2$.

$$\ln L\left(\xi,\sigma \mid \underline{x}\right) = \begin{array}{l} -m\ln\sigma - \left(\dfrac{1}{\xi}+1\right)\displaystyle\sum_{i=1}^{m}\ln\left(1+\dfrac{\xi}{\sigma}\left(x_i-\tau\right)\right) \\ -m\ln\left(1-\left(1+\dfrac{\xi\left(T-\tau\right)}{\sigma}\right)^{-\frac{1}{\xi}}\right) \end{array}, \quad \sigma \neq 0 . \qquad (3.4.1)$$

Hence it is difficult to simulate the posterior distribution using standard techniques (e.g. inverse transformation). There are, however, techniques for sampling from non-standard distributions, such as approximation with the importance sampling method [13] or the exact Gibbs/ARMS sampling algorithm [10],[9].

The Gibbs/ARMS algorithm is a combination of the Gibbs sampling method [10] for sampling a multivariate distribution from univariate conditional distributions, and the ARMS (Adaptive Rejection Metropolis Sampling) algorithm [9] for sampling from a general univariate distribution given its PDF. Together, the Gibbs/ARMS algorithm enables sampling from a general multivariate distribution by specifying the probability density function - which with uniform priors is simply the likelihood.

The Gibbs/ARMS algorithm retains the dependency between parameters, as can be seen from the posterior joint distribution density plot in Figure 1.

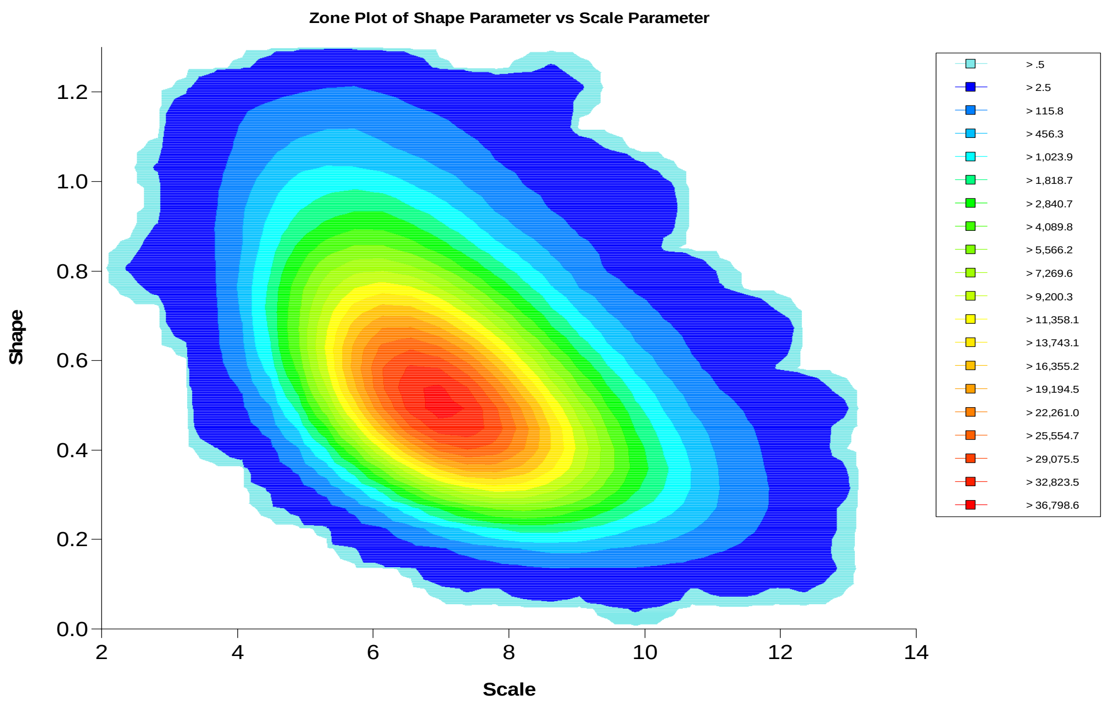

Figure 1: Simulated bivariate distribution of scale and shape parameters of the GPD, based on 100,000 simulations. The colours indicate the density of generated samples in a small region around that point.

In Table 4, we compare some statistics of the simulated distribution of parameters with the asymptotic approximation of the distribution of maximum likelihood estimators. We note that when the posterior distribution is skewed, we do not expect the mean of the posterior distribution to correspond to the maximum likelihood estimate. In fact, the maximum likelihood estimate corresponds to the mode of the posterior distribution.

Table 4: Comparison of mean and standard deviation for severity parameters using asymptotic distribution of MLE and Bayesian approach.

	Bayes	Asymptotic MLE
Estimate [scale]	7.080	6.960
Estimate [shape]	0.546	0.502
Standard Error [scale]	1.143	1.117
Standard Error [shape]	0.153	0.141
Correlation [scale,shape]	-0.511	-0.543

4. Forecasting, Capital Requirements & Reinsurance Pricing

4.1 Aggregate Distributions With and Without Parameter Uncertainty.

In this section we present the results of 100,000 Monte-Carlo simulations in the simulation engine, Igloo Professional, of the aggregate claims distribution, using both fixed parameters and parameters simulated from the posterior distribution using the parameterisation tool ExtrEMB.

To include parameter uncertainty, we use the predictive distributions for both the frequency and severity (as described in section 3.2) to model the losses. The predictive distribution for the number of losses is simulated by first simulating the Poisson parameter μ from the gamma distribution as defined in section 3.3, and then for each simulation drawing a single number from a Poisson distribution with that parameter to represent the number of claims. For each simulation of claim numbers, parameter couples are drawn from the bivariate posterior distribution of the severity parameters using the Gibbs/ARMS algorithm. The individual claim amounts are then simulated from the truncated generalised Pareto distribution with the simulated parameters. For comparison, we simulate the predictive distribution using fixed maximum likelihood parameters. Figure 2 shows the upper tail of the resultant aggregate claims distribution both including parameter uncertainty and using fixed parameters.

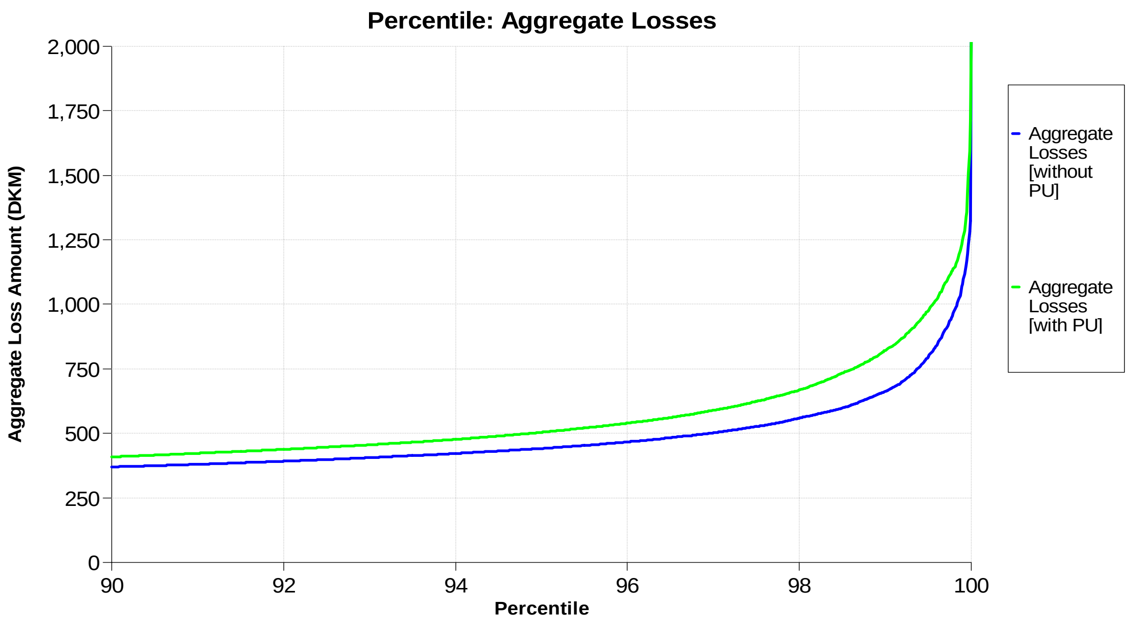

Figure 2: Percentile graph above the 90th percentile of aggregate distribution with parameter uncertainty (green line) and without parameter uncertainty (blue line), based on 100,000 simulations.

Some summary statistics of the predictive distributions are shown in Table 5.

Table 5: Summary statistics of aggregate distributions

	Aggregate Distribution without Parameter Uncertainty (DKM)	Aggregate Distribution with Parameter Uncertainty (DKM)
Mean	233	250
St. Dev	121	147
Median	210	219
90th percentile	372	411
95th percentile	445	509
99th percentile	676	828
99.5th percentile	800	987

We note that, when including parameter uncertainty, the mean of the aggregate distribution is increased. This effect is also seen in the frequency and severity distributions.

4.2 Capital Requirements

In section 4.1, we have simulated the aggregate loss distribution. In this section, we will use the simulated loss distribution to assess the amount of capital required by the insurer to support this single line of business. To measure capital requirements, we use the underwriting profit, defined as premium less claims, as a risk profile. We assume a gross premium of 300 DKM. Figure 3 shows the gross underwriting profit distribution for this example.

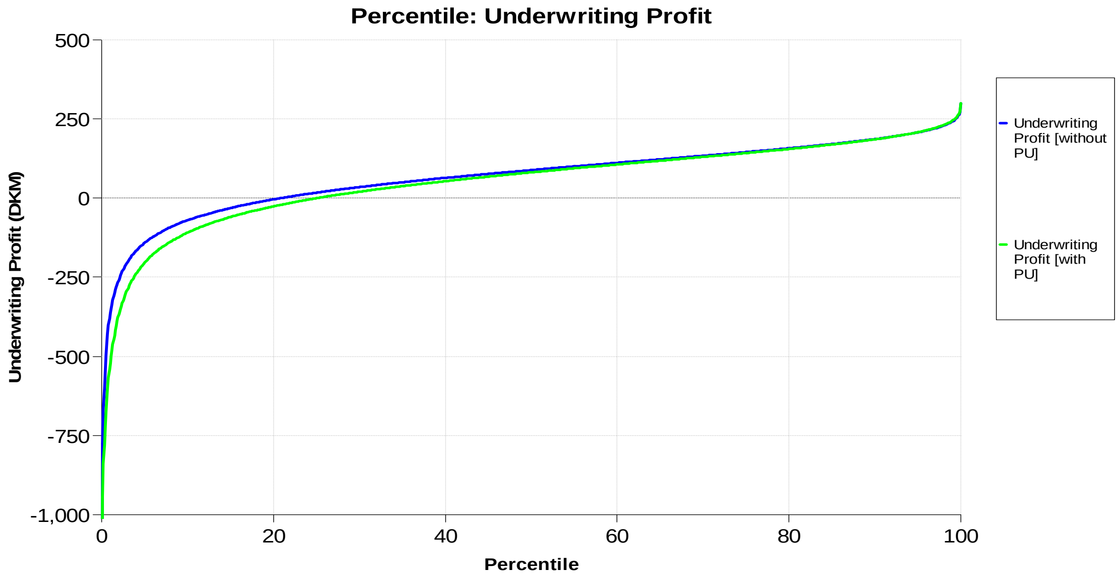

Figure 3: Gross underwriting profit distributions with and without parameter uncertainty.

To derive the capital requirements, we have used two risk measures, VaR and TVaR. For VaR, we set the risk criteria at 99.5% and for TVaR at 98.7%, such that the capital required under both measures is the same. We apply these risk measures to the risk profiles both with and without parameter uncertainty. The results can be seen in Table 6.

Table 6: Capital requirements

Method	Without Parameter Uncertainty (DKM)	With Parameter Uncertainty (DKM)
99.5% VaR Capital Requirement (DKM)	500	687
98.7% TVaR Capital Requirement (DKM)	500	676

4.3 Reinsurance Pricing

Now, we look at the situation faced by a reinsurer when pricing a reinsurance contract based on the same claims distribution. As an example, a single-layer individual excess of loss contract is investigated. The reinsurance contract characteristics are detailed in Table 7. The characteristics have been chosen to mimic a typical market contract.

Table 7: Characteristics of the reinsurance contract (all in DKM)

	Limit	Excess	Aggregate Deductible	Reinstatements
Layer 1	80	40	80	2 free

Figure 4 shows the upper tail of the cumulative distribution function of the aggregate reinsurance recoveries under the reinsurance contract. Table 8 shows the mean and standard deviation of the aggregate recoveries.

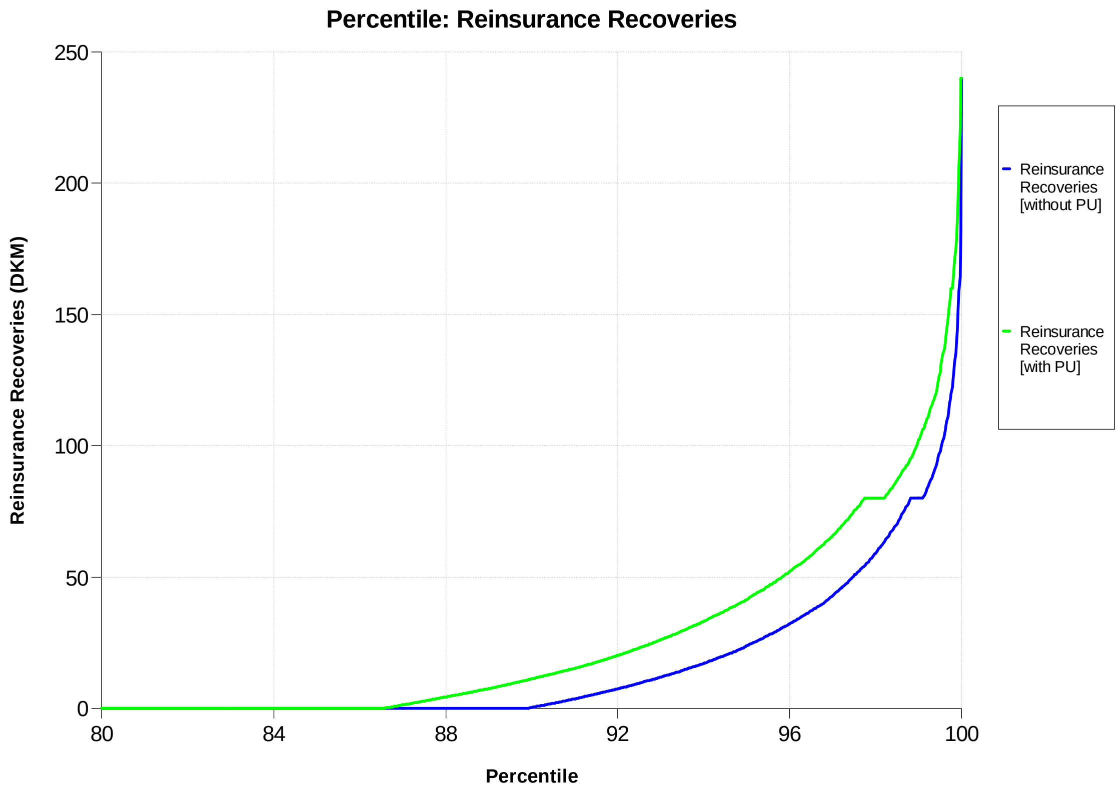

Figure 4: Cumulative distribution function of aggregate reinsurance recoveries with parameter uncertainty (green line) and without parameter uncertainty (blue line).

Table 8: Average and standard deviation of aggregate reinsurance recoveries, together with a price for the reinsurance based on 25% standard deviation loading.

Method	Without Parameter Uncertainty (DKM)	With Parameter Uncertainty (DKM)
Average RI Recovery (Million DK)	3.5	5.6
St. Dev. RI Recovery (Million DK)	14.5	20.5
Reinsurance Price with 25% SD loading	7.1	10.8

For the whole range of simulations, the reinsurance recoveries from this contract are always greater and more volatile when considering parameter uncertainty. Providing that a rational pricing technique is used, the price of the reinsurance will be greater when considering parameter uncertainty. As an example, a reinsurer pricing this contract using the mean recovery with 25% standard deviation loading and ignoring parameter uncertainty will possibly misestimate the risk in this contract.

4.4 Summary of Results

We have applied the classical MLE approach to fit a frequency-severity model for large claims with an upper-truncated severity distribution using point parameter estimates. We have then used the Bayesian approach to incorporate parameter uncertainty into this model. Monte Carlo techniques were used to generate 100,000 simulations of the aggregate predictive claims distribution, including parameter uncertainty. This served to assess capital requirements using both VaR and TVaR at different risk appetites and to calculate recoveries and price of an example of a reinsurance contract.

Including parameter uncertainty increases capital requirements by approximately 36% for the example investigated. The average reinsurance recovery on an example excess of loss contract is increased by nearly 60%. The standard deviation of reinsurance recoveries also increases. For most pricing methods, parameter uncertainty will increase the estimated price of this contract.

5. Conclusion

We have investigated the effect of parameter uncertainty in the frequency-severity model from both a capital adequacy and reinsurance pricing perspective. The effect of parameter uncertainty is a significant increase in both capital required and reinsurance recoveries. The inclusion of parameter uncertainty acts to increase the mean of the aggregate predictive claims distribution. This may strengthen the case for greater prudence by actuaries when assessing future claims experience.

An interesting topic for future research would be an analysis of the impact of other sources of uncertainty using the Bayesian framework. Model uncertainty can arise through the choice of distribution or the choice of thresholds for the selected distribution. Another source of uncertainty is data uncertainty, which can arise when adjustments are made to raw data, such as IBNER, IBNR, trending or inflation. In the data set of Danish fire losses used in this paper, we have not performed any data adjustments, so as to be consistent with other authors.

The effect of parameter uncertainty is not limited to large claims: the same methodology can be applied to any other aspect of DFA modelling, such as attritional claims or even dependency structures. Parameter uncertainty is usually taken into account in modelling reserving risk via bootstrapping or Bayesian methods [8]. The results of this study show that parameter uncertainty can have a significant impact on both capital requirements and reinsurance pricing, and should not be ignored.

Acknowledgements

The authors would like to thank their colleagues at EMB Consultancy, in particular Dr. Peter England for valuable comments on an earlier draft of this paper. Thanks also go to Mette Havning for providing the sample data set.

References

[1] Bohning, D. 1994. A note on a test for Poisson overdispersion. *Biometrika*. 81:418-419.

[2] Box, G. E. P. and G. C. Tiao. 1992. *Bayesian inference in statistical analysis*. Wiley.

[3] Cairns, A. J. 2000. A discussion of parameter and model uncertainty in insurance. *Insurance: Mathematics and Economics* 27:313–330.

[4] D'Agostino, R. B. and M. A. Stephens. *Goodness-of-Fit Techniques,* Marcel Dekker, Inc.

[5] Efron, B. 1979. Bootstrap Methods: Another Look at the Jackknife. *Annals of Statistics.* Vol 7 pp. 1-26.

[6] Efron, B., and R. J. Tibshirani. 1993. *An Introduction to the Bootstrap*. Chapman and Hall.

[7] Embrechts P., C. Klueppelberg and T. Mikosch. 1997. *Modelling Extremal Events for Insurance and Finance.* Springer.

[8] England P. D. and R. J. Verrall. 2002. Stochastic Claims Reserving in General Insurance. *British Actuarial Journal*. Volume 8, Number 3, pp. 443-518(76).

[9] Gilks, W. R., N. G. Best and K. K. C. Tan. 1994. Adaptive Rejection Metropolis Sampling within Gibbs Sampling. *Applied Statistics* 44:455–472.

[10] Gilks, W. R., S.Richardson and D. J. Spiegelhalter. 1995. *Markov Chain Monte Carlo in Practice*. Chapman and Hall.

[11] Klugman, S. A., H. H. Panjer and G. E. Willmot. 1998. *Loss Models: From Data to Decisions.* Wiley.

[12] Kreps, R. E. 1997. Parameter Uncertainty in (Log) Normal Distributions. *Proceedings of the Casualty Actuarial Society*. Vol: LXXXIV pp. 553-580.

[13] Lee, P. M. 1997. *Bayesian Statistics: An Introduction*. 3rd ed. Arnold.

[14] Mata, A. 2000. Parameter Uncertainty For Extreme Value Distributions. *GIRO Convention Papers*.

[15] McNeil, A. J. 1997. Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory, *ASTIN Bulletin, International Actuarial Association.* Vol: 27:1 pp. 117-137.

[16] Meyers, G. G. and N. Schenker. 1982. Parameter Uncertainty in the Collective Risk Model. *Casualty Actuarial Society Discussion Paper Program.* Vol: May pp. 253-300 .

[17] Millns, R. and R. Weston. 2005. Parameter Uncertainty and Capital Modelling. *GIRO Convention.*

[18] Resnick, S. I. 1999. Discussion of the Danish Data on Large Fire Insurance Losses. *ASTIN Bulletin, International Actuarial Association* Vol: 27:1 pp: 139-151.

[19] Rytgaard, M. 1990. Estimation in the Pareto Distribution. *ASTIN Bulletin, International Actuarial Association* Vol: 20:2 pp. 201-216.