Bayesian Modelling of Financial Guarantee Insurance

Anne Puustelli [1]

Department of Mathematics, Statistics and Philosophy, Statistics Unit, FIN-33014 University of Tampere, Finland

Lasse Koskinen [1]

Insurance Supervisory Authority of Finland, P.O. Box 449, FIN-00101 Helsinki, Finland
and
Helsinki School of Economics, P.O. Box 1210, FIN-00101 Helsinki, Finland

Arto Luoma [*,1]

Department of Mathematics, Statistics and Philosophy, Statistics Unit, FIN-33014 University of Tampere, Finland.

Abstract

In this research we model the claim process of financial guarantee insurance and predict the pure premium and the required amount of risk capital. The used data is from the financial guarantee system of the Finnish statutory pension scheme. The losses in financial guarantee insurance may be devastating during an economic depression (that is, deep recession). This indicates that the economic business cycle, and in particular depressions, must be taken into account when the claim amounts of financial guarantee insurance are modelled. A Markov regime-switching model is used to predict the number and length of depression periods in the future. The claim amounts are predicted by using a transfer function model where the predicted growth rate of real GNP is an explanatory variable. The pure premium and initial risk reserve are evaluated on the basis of the predictive distribution of claim amounts. Bayesian methods are applied throughout the modelling process. For example, the Gibbs sampler is used in the estimation of the business cycle model. Simulation results show that the required amount of risk capital is high even though depression is an infrequent phenomenon.

Key words: Business cycle, Gibbs sampler, Hamilton model, Risk capital, Surety insurance
PACS: C11, G22, G32, IM22, IM41

1 Introduction

A guarantee insurance (surety insurance) is typically required when there is a doubt of the fulfilment of a contractual, legal or regulatory obligation. It is designed to protect some public or private interest from the consequences of the default or delinquency of another party. Financial guarantee insurance covers losses from specific financial transactions. Due to differences in laws and regulations guarantee insurance is a country specific business; see, for example, Sigma (2006).

When a country experiences economic depression (that is, deep recession) the losses in financial guarantee insurance may reach catastrophic dimensions for several years. During that time the number of claims can be extraordinary large and, what is more important, the proportion of excessive claims can be much higher than in usual periods (see, for example, Romppainen, 1996; Sigma, 2006). As the future growth of the economy is uncertain, it is important to consider the level of uncertainty one can expect in the future claim process. A mild and short downturn in the national economy increases the losses suffered by financial guarantee insurers only moderately, but severe downturns in the national economy are crucial. History knows several economic depressions. These include the Great Depression in the 1930s, World Wars I and II, and the oil crisis in the 1970s. In recent years the Finnish experience from the beginning of the 1990s and the Asian crisis in the late 1990s are good examples. An interesting statistical approach for analyzing the timing and effects of the Great Depression is the Regime Switching method in Coe (2002).

Here we model the claim process of financial guarantee insurance in the economic business cycle context. We build on the following three studies on the financial guarantee system of the Finnish pension scheme. Rantala and Hietikko (1988) modelled the solvency issues by means of linear models. Their main objective was to test methods for specifying bounds for the solvency capital. The linear method combined with the data not containing any fatal depression period - Finland was struck by depression in the early 1990s after

* Corresponding author, telephone: +358-3-35516428, fax:+358-3-35516157
Email addresses: anne.puustelli@uta.fi (Anne Puustelli), lasse.koskinen@vakuutusvalvonta.fi (Lasse Koskinen), arto.luoma@uta.fi (Arto Luoma).
[1] The authors of this study are grateful to Yrjö Romppainen for the data he provided us, the most valuable comments and inspiring discussions. We also would like to thank Vesa Ronkainen for his useful comments.

the article was published - underestimated risk. Romppainen (1996) analyzed the structure of the claim process during the depression period. Koskinen and Pukkila (2002) also applied the economic cycle model. Their simple model gives approximate results but lacks sound statistical ground. We use modern statistical methods that offer advantages for assessing the uncertainty.

There is no single "best practice" model for credit risk capital assessment (Alexander, 2005). The main approaches are structural firm-value models, option-theoretical approaches, rating-based methods, macroeconomic models and actuarial loss models. In contrast to market risk, there has been little detailed analysis of the empirical merits of different models. A review of commonly used financial mathematics methods can be found for example in McNeil et al. (2005). Since the guarantee loans are nontraded in Finland, we adopt an actuarial approach. However, this approach is also difficult, since depression is an exceptional event and its modelling is difficult with standard actuarial methods.

Cairns (2000) points out that uncertainty in actuarial modelling arises from three sources: 1) uncertainty due to the stochastic nature of a given model; 2) uncertainty in the values of the parameters in a given model; 3) uncertainty in the model underlying what we are able to observe and determining the quantity of interest. In financial guarantee insurance the main problem seems to fall into the third category because of the complexity of the underlying risk process.

From the methodological point of view we adopt the Bayesian approach recommended for example by Scollnik (2001). Simplified models or simplified assumptions may fail to reveal the true magnitude of the risks faced by the insurer. While undue complexity is generally undesirable, there may be situations where complexity cannot be avoided. Best et al. (1996) explain how Bayesian analysis can generally be used for realistically complex models. An example of concrete modelling is provided by Hardy (2002), who applies Bayesian techniques to a regime-switching model of the stock price process for risk management purposes. Another example can be found in Smith and Goodman (2000), who present models for the extreme values of large claims and use modern techniques of Bayesian inference. Here, Bayesian methods are used throughout the modelling process, for example the Gibbs sampler in the estimation of the business cycle model. The proposed actuarial model is used for simulating purposes in order to study the effect of the economic cycle on the needed pure premium and initial risk reserve.

We apply the Markov regime-switching model to predict the number and length of depression periods in the future. The prediction of claim amounts is made by using a transfer function model where the predicted growth rate of real GNP is an explanatory variable. More specifically, we utilize the business cycle model introduced by Hamilton (1989). In the Hamilton method all the dating decisions or, more correctly, the probability that a particular time period is in recession, are based on the observed data. The method assumes that there are two distinct states (regimes) in the business cycle - one for expansion and one for recession - that are governed by a Markov chain. The stochastic nature of the GNP growth depends on the prevailing state.

The financial guarantee insurance is characterized by long periods of low loss activity punctuated by short severe spikes; see Sigma (2006). As such conventional dichotomic business models are inadequate, since severe recessions constitute the real risk. We propose a model where the two states represent 1) the depression period state and 2) its complement state consisting of both boom and mild recession periods. We use Finnish real GNP data to estimate our model. The claim data is from the financial guarantee insurance system of the Finnish pension scheme. Combining a suitable business cycle model with a transfer function model provides a new way to analyze the solvency of a financial guarantee provider with respect to claim risk.

The paper is organized as follows. In Section 2 the Finnish credit crisis in the 1990s is described. Section 3 introduces the business cycle model and Section 4 presents the transfer function model and predictions. Model checks are presented in Section 5. Section 6 concludes.

2 The Finnish experience in the 1990s

During the years $1991-1993$ Finland's GNP dropped by 12%. Naturally that period was harmful to all sectors of the economy and society as a whole. However, the injuries suffered in the insurance sector were only moderate, at least compared with the problems of the banking sector at the same time. An important exception was the financial guarantee insurance related to the statutory earnings-related pension scheme of the private sector. At a general level, Norberg (2006) describes the risk presented to pension schemes under economic and demographic developments.

The administration of the pension scheme is decentralised to numerous insurance companies, company pension funds and industry-wide pension funds. The central body for the pension scheme is the Finnish Centre for Pensions (FCfP). The special financial guarantee insurance was administrated by the FCfP. It was designed to be a guarantee for loans granted by pension insurance companies as well as to secure the assets of pension funds.

The business was started in 1962 and continued successfully until Finland was hit by depression in the 1990s. Then the losses took catastrophic dimensions

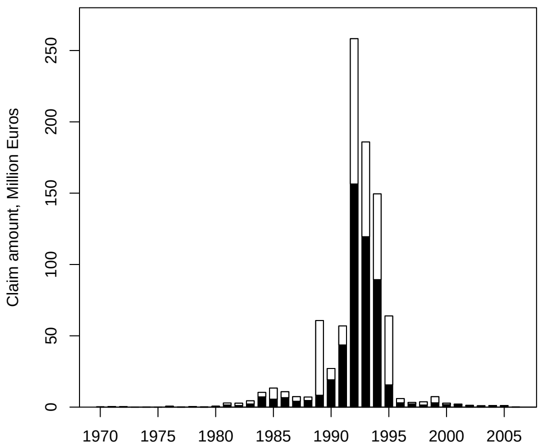

Fig. 1. Claims paid from financial guarantee insurance by the Finnish Centre for Pensions of Finland between 1980 and 2000. The lower dark part of the bar describes the final loss by August 2001.

and the financial guarantee insurance activity of the FCfP was closed. Claims paid by the FCfP are shown in Figure 1. The cost was levied to all employers involved in the mandatory scheme. Hence, pension benefits were not jeopardized.

Afterwards the FCfP's run-off portfolio was transferred to the new company named "Garantia". A more detailed description of the case of the FCfP can be found in Romppainen (1996). In order to promote the capital supply, the FCfP had a legal obligation to grant financial guarantee insurance to company pension funds and industry-wide pension funds for which insurance was obligatory. Hence, it employed fairly liberal risk selection and tariffs. This probably had an influence on the magnitude of the losses. Hence, the data reported by Romppainen and used here can not be expected, as such, to be applicable in other environments. The risks would be smaller in conventional financial guarantee insurance, which operates solely on a commercial basis.

It is interesting to note that there are similar problems also in the USA at present. The corresponding US institute is Pension Benefit Guaranty Corporation (PBGC). It is a federal corporation created by the Employee Retirement Income Security Act of 1974. It currently protects the pensions of nearly 44 million American workers and retirees in 30,330 private single-employer and multiemployer defined benefit pension plans. Pension Insurance Data Book 2005 (page 31) reveals that total claims of PBGC have increased rapidly from about 100 million dollars in 2000 to 10.8 billion dollars in 2005. This increase in claims can not be explained by nation-wide depression, but it may be related to problems of special industry sectors (for example aviation).

3 National economic business cycle model

Our first goal is to find a model by which we can forecast the growth rate of GNP. We will use annual Finnish data on real GNP from 1860 to 2004, provided by Statistics Finland. We are particularly interested in the number and length depression periods. For this purpose we will utilize the Markov regime-switching model introduced by Hamilton (1989). The original Hamilton model has two states for the business cycle: expansion and recession. In our situation, however, it is more important to detect depression, since it is the period when financial guarantee insurance will suffer its most severe losses. Therefore, we will define the states in a slightly different way in our application. Specifically, we will use a two-state regime-switching model in which the first state covers both expansion and recession periods and the second state is for depression.

Our estimation results correspond to this new definition, since depression periods are included in our data set. By contrast, Hamilton used quarterly U.S. data from 1951 to 1984, which do not include years of depression.

The Hamilton model may be expressed as $y_t = \alpha_0 + \alpha_1 s_t + z_t$, where y_t denotes the growth rate of real GNP at time t, s_t the state of the economy and z_t a zero-mean stationary random process, independent of s_t. The parameters α_0 and α_1 and the state s_t are unobservable and should be estimated. We will assume that z_t is an autoregressive process of order r, denoted by $z_t \sim \mathrm{AR}(r)$. It is defined by the equation $z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_r z_{t-r} + \epsilon_t$, where $\epsilon_t \sim \mathrm{N}(0, \sigma_\epsilon^2)$ is an i.i.d. Gaussian error process. The growth rate at time t is calculated as $y_t = \log(\mathrm{GNP}_t) - \log(\mathrm{GNP}_{t-1})$.

We define the state variable s_t to be 0, when the economy is in expansion or recession, and 1, when it is in depression. The transitions between the states are controlled by the first-order Markov process with transition probabilities

$$
\begin{aligned}
\mathrm{P}(s_{t+1} = 0 | s_t = 0) &= p, \\
\mathrm{P}(s_{t+1} = 1 | s_t = 0) &= 1 - p, \\
\mathrm{P}(s_{t+1} = 0 | s_t = 1) &= 1 - q, \\
\mathrm{P}(s_{t+1} = 1 | s_t = 1) &= q.
\end{aligned}
$$

Thus, the transition matrix is given by

$$
P = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}.
$$

The stationary probabilities $\boldsymbol{\pi} = (\pi_0, \pi)'$ of the Markov chain satisfy the equations $\boldsymbol{\pi}' P = \boldsymbol{\pi}'$ and $\boldsymbol{\pi}'\mathbf{1} = 1$, where $\mathbf{1} = (1,1)'$.

The Hamilton model was originally estimated by maximising the marginal likelihood of the observed data series y_t. Then the probabilities of the states were calculated conditional on these maximum likelihood estimates. The numerical evaluation was done by a kind of nonlinear version of the Kalman filter. By contrast, we will use Bayesian computation techniques throughout this paper. Their advantage is that we need not rely on asymptotic inference and the inference on the state variables is not conditional on the parameter estimates. The Hamilton model will be estimated using the Gibbs sampler, introduced by Geman and Geman (1984) in the context of image restoration. Examples of Gibbs sampling can be found in Gelfand et al. (1990) and Gelman et al. (2004). Carlin et al. (1992) provide a general approach to its use in nonlinear state-space modelling.

The Gibbs sampler, also called alternating conditional sampling, is a useful algorithm for simulating multivariate Markov chains. It is defined in terms of subvectors of $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_p)$, where $\boldsymbol{\theta}$ is the random vector whose distribution is to be simulated. In each iteration the Gibbs sampler goes through $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_p$ and draws values from their conditional distributions, conditional on the latest values of the other components of $\boldsymbol{\theta}$. It can be shown that this algorithm produces an ergodic Markov chain whose stationary distribution is the distribution of $\boldsymbol{\theta}$. In order to use the Gibbs sampler as an estimating algorithm, all full conditional posterior distributions of the parameters need to be evaluated.

To simplify some of the expressions we will use the following notations: $\mathbf{y} = (y_1, y_2, \ldots, y_T)'$, $\mathbf{s} = (s_1, s_2, \ldots, s_T)'$ and $\mathbf{z} = (z_1, z_2, \ldots, z_T)'$. We will also need the matrix

$$\mathbf{Z} = \begin{pmatrix} \mathbf{z}_0' \\ \mathbf{z}_1' \\ \vdots \\ \mathbf{z}_{T-1}' \end{pmatrix},$$

whose rows are of the form $\mathbf{z_t} = (z_t, z_{t-1} \ldots, z_{t-r+1})'$. Furthermore, we denote the vector of autoregressive coefficients by $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_r)'$ and the vector of all parameters by $\boldsymbol{\eta} = (\alpha_0, \alpha_1, \boldsymbol{\phi}', \sigma_\epsilon^2, p, q)'$. In the following treatment we assume the pre-sample values $\mathbf{y}_0 = (y_0, ..., y_{1-r})'$ and $\mathbf{s}_0 = (s_0, ..., s_{1-r})'$ to be known. In fact, $\mathbf{s}_0$ is not known, but we will simulate its components in a similar way as those of $\mathbf{s}$.

Using these notations the density of $\mathbf{y}$, conditional on $\mathbf{s}$ and the parameters $\boldsymbol{\eta}$, can be written as

$$p(\mathbf{y}|\mathbf{s}, \boldsymbol{\eta}) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} \exp\left(-\frac{1}{2\sigma_\epsilon^2}(y_t - \alpha_0 - \alpha_1 s_t - \boldsymbol{\phi}'\mathbf{z}_{t-1})^2\right).$$

In order to make computations easy, we use the following prior distributions:

$$\begin{aligned}
&p \sim \text{Beta}(\alpha_p, \beta_p), \\
&q \sim \text{Beta}(\alpha_q, \beta_q), \\
&p(\boldsymbol{\phi}, \sigma_\epsilon^2) \propto \frac{1}{\sigma_\epsilon^2}, \\
&p(\alpha_0) \propto 1, \\
&p(\alpha_1) \propto \text{N}(\alpha_1|\mu_0, \sigma_0^2) \times I(\alpha_1 < -0.03).
\end{aligned}$$

We obtained noninformative prior distributions for p and q by specifying as prior parameters $\alpha_p = \beta_p = \alpha_q = \beta_q = 0.5$. These values correspond to Jeffreys uninformative prior distribution in the standard Bernoulli model. We also made sensitivity analysis by using informative priors with parameters $\alpha_p = 19$, $\beta_p = 3$; $\alpha_q = \beta_q = 11$. These values correspond to the idea that the chain has switched to state 1 in 2 out of 20 prior cases when it has been is state 0 and has switched to state 0 in 10 out of 20 prior cases when it has been in state 1. By this choice of priors we could increase the probability of the state 1 so that it corresponded to our concept of depression.

For α_0, $\boldsymbol{\phi}$ and σ_ϵ^2 we gave improper, noninformative prior distributions. The prior distribution of α_1 prevents it from getting a positive value (that is, the state can then be interpreted as a depression state). Here, notation $\text{N}(\alpha_1|\mu_0, \sigma_0^2)$ refers to the Gaussian density with mean μ_0 and variance σ_0^2 and $I(\alpha_1 < -0.03)$ the indicator function obtaining the value 1, if $\alpha_1 < -0.03$, and 0, otherwise. We specified the values $\mu_0 = -0.1$, $\sigma_0^2 = 0.2^2$ as prior parameters, which results in a fairly noninformative prior distribution. We also experimented here with an informative alternative $\mu_0 = -0.05$, $\sigma_0^2 = 0.025^2$, which reduces the difference between the states and increases the probability of state 1. The results were similar to those obtained when informative priors were given to p and q.

In order to implement the Gibbs sampler the full conditional posterior distributions of the parameters are needed:

$$\{p|\mathbf{s}\} \sim \text{Beta}\left(\sum_{t=1}^{T}[(1-s_t)(1-s_{t-1})] + \alpha_p, \sum_{t=1}^{T}[s_t(1-s_{t-1})] + \beta_p\right),$$

$$\{q|\mathbf{s}\} \sim \text{Beta}\left(\sum_{t=1}^{T}[s_t s_{t-1}] + \alpha_q, \sum_{t=1}^{T}[s_{t-1}(1-s_t)] + \beta_q\right),$$

$$\{s_t|\mathbf{s}_{(-t)}, \mathbf{y}, \boldsymbol{\eta}\} \\ \sim \text{Bernoulli}\left(\frac{\Pr(s_t = 1|\mathbf{s}_{(-t)}, \boldsymbol{\eta}) p(\mathbf{y}|s_t = 1, \mathbf{s}_{(-t)}, \boldsymbol{\eta})}{\sum_{j=0}^{1} \Pr(s_t = j|\mathbf{s}_{(-t)}, \boldsymbol{\eta}) p(\mathbf{y}|s_t = j, \mathbf{s}_{(-t)}, \boldsymbol{\eta})}\right), t = 1, \ldots, T,$$

$$\{\boldsymbol{\phi}|\mathbf{y}, \mathbf{s}, \alpha_0, \alpha_1, \sigma_\epsilon^2\} \sim \text{N}\left((\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{z}, \sigma_\epsilon^2(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\right),$$

$$\{\sigma_\epsilon^2|\mathbf{y}, \mathbf{s}, \alpha_0, \alpha_1, \boldsymbol{\phi}\} \sim \text{Inv-}\chi^2\left(T, (\boldsymbol{z} - \boldsymbol{Z}\boldsymbol{\phi})'(\boldsymbol{z} - \boldsymbol{Z}\boldsymbol{\phi})/T\right),$$

$$\{\alpha_0|\mathbf{y}, \mathbf{s}, \alpha_1, \boldsymbol{\phi}, \sigma_\epsilon\} \sim \text{N}\left(\frac{\sum_{t=1}^{T} y_t^*}{T}, \frac{\sigma_\epsilon^2}{T}\right),$$

$$p(\alpha_1|\mathbf{y}, \mathbf{s}, \alpha_0, \boldsymbol{\phi}, \sigma_\epsilon) \propto \text{N}(\alpha_1|\hat{\alpha}_1, \hat{\sigma_1}^2) \times I(\alpha_1 < -0.03),$$

where we have denoted

$$\hat{\alpha}_1 = \frac{\frac{\sum_{t=1}^{T} s_t y_t^{**}}{\sigma_\epsilon^2} + \frac{1}{\sigma_0^2}\mu_0}{\frac{\sum_{t=1}^{T} s_t}{\sigma_\epsilon^2} + \frac{1}{\sigma_0^2}}, \qquad \hat{\sigma}_1^2 = \left(\frac{\sum_{t=1}^{T} s_t}{\sigma_\epsilon^2} + \frac{1}{\sigma_0^2}\right)^{-1},$$

and

$$y_t^* = y_t - \alpha_1 s_t - \boldsymbol{\phi}'\mathbf{z}_{t-1}, \qquad y_t^{**} = y_t - \alpha_0 - \boldsymbol{\phi}'\mathbf{z}_{t-1}.$$

The notation Inv-$\chi^2(\nu, s^2)$ means the scaled inverse-chi-square distribution, defined as $\frac{\nu s^2}{\chi_\nu^2}$, where χ_ν^2 is a chi-square distributed random variable with ν degrees of freedom.

Note that the probability of state 1 at time t, given the states at the other time points, $\mathbf{s}_{(-t)}$, is easily calculated as

$$\Pr(s_t = 1|\mathbf{s}_{(-t)}, \boldsymbol{\eta}) = \frac{\Pr(s_t = 1|s_{t-1}, p, q)\Pr(s_{t+1}|s_t = 1, p, q)}{\Pr(s_{t+1}|s_{t-1}, p, q)}, \quad 0 < t < T.$$

In Figure 2, one simulated chain, produced by the Gibbs sampler, is shown. As we can see, the chain converges rapidly to its stationary distribution and the component series of the chain mix well, that is, they are not too autocorrelated. The summary of the estimation results, based on three simulated chains, as well as Gelman and Rubin's diagnostics (Gelman et al., 2004) are given in Appendix A. The values of the diagnostic are close to 1 and thus indicate good convergence.

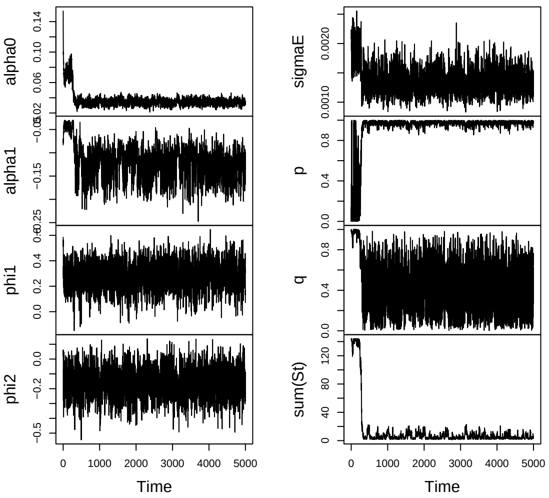

Fig. 2. Iterations of the Gibbs sampler.

4 Prediction of the claim amounts

Our ultimate goal is to predict the pure premium and the required amount of risk capital needed for the claim deviation. The claim data is obtained from FCfP (see Section 2) and the years included in this study are $1966-2004$. The prediction of the claim amounts is made by using a transfer function model of the form

$$x_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 y_t + \epsilon_t,$$

where $x_t = \Phi^{-1}(x_t^*)$ and x_t^* is the proportion of gross claim amount to technical provision at time t, y_t the growth rate of GNP and $\epsilon_t \sim \mathrm{N}(0, \sigma^2)$ an i.i.d. Gaussian error process. The parameters β_0, β_1 and β_2 are not known and are estimated. The series y_t is predicted using the model described in the previous section.

The model $x_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 y_t + \epsilon_t$ may also be expressed in the form

$$\begin{aligned} x_t - \beta_0 &= \frac{\beta_2}{1-\beta_1 B} y_t + \frac{1}{1-\beta_1 B}\epsilon_t \\ &= \beta_2(y_t + \beta_1 y_{t-1} + \beta_1^2 y_{t-2} + \dots) + \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_1^2 \epsilon_{t-2} + \dots , \end{aligned}$$

from which one can see that x_t can be obtained by applying exponential filters to the input series y_t and ϵ_t.

In our analysis we applied the probit transformation $\Phi^{-1}(.)$, the inverse function of the standard normal distribution function, to the variable x_t^*. This is reasonable, since x_t^* cannot exceed 1. As an alternative we used the logit link

$x_t = \text{logit}(x_t^*) = \log(x_t^*/(1 - x_t^*))$. The parameters β_0, β_1 and β_2 were estimated using standard Bayesian simulation for regression models. We used an otherwise uninformative prior distribution but made the restriction $\beta_1 < 1$ to ensure that that the estimated model for x_t is stationary.

The premium and the initial risk reserve are evaluated from the posterior predictive distribution of the proportions of claim amount to technical provision. For simplicity the technical provision is set at 1 in the prediction. Then the predicted proportions are the same as the predicted claim amounts. The premium is set at the overall mean of all iterations through the prediction period. The balance at time t is calculated by subtracting the claim amount at time t from the cumulated premiums. The five-year 95% and 75% values at risk are evaluated from the minimum balance of each iteration. The distribution of the minimum balance is extremely skewed, which can be explained by the rareness of depression and by the huge losses of guarantee insurance once depression hits. This phenomenon can be seen from Figure 3, which shows the 95% prediction limits of the balance when both probit and logit links are used. In both cases, noninformative prior distributions were used in the estimation of the Hamilton model. The curves indicating the lower prediction limits are much steeper than those indicating the upper limits. The use of the logit link produces more extreme simulation paths.

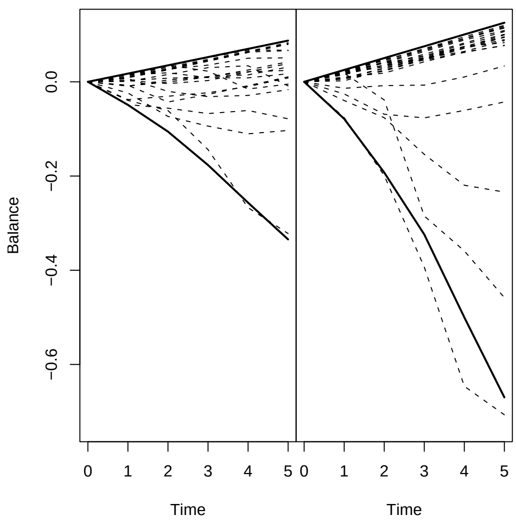

Fig. 3. Simulation results for the balance of the guarantee insurance. The solid lines indicate the 95% values at risk and the dashed lines 20 example paths. The simulation results based on the probit and logit links are shown in the left and right parts of the figure, respectively.

Extensive simulations (5 000 000 iterations) were done to evaluate the pure premium and the 95% and 75% values at risk for the prediction period of five years. The pure premium levels were 1.9% and 2.7% with the probit and logit links, respecitvely. The 95% values at risk were 2.3 and 2.9 times the five-year premium with the probit and logit links, respectively. The corresponding figures for the 75 % value at risk were 0.30 and 0.18, respectively. These results were obtained when noninformative prior distributions were used in the estimation of the Hamilton model. When informative prior distributions were used, the results did not change considerably. The choice between the probit and logit links in the context of the transfer function model had a larger effect. When the logit link was used, the estimate of the premium level became unrealistically large.

5 Model checks for the Hamilton model and the transfer function model

We made some sensitivity analysis with respect to the prior distributions related to the Hamilton model. We found that by using informative prior distributions for the transition probabilities p and q we could increase the estimated probabilities of state 1 so that it corresponded better to its interpretation as depression. This can be seen from Figure 4 where the growth rate of GNP is shown along with the probabilities of depression, estimated using two different kinds of prior information. The same goal was achieved by giving an informative prior for α_1. However, these adjustments did not considerably effect the estimated premium or value at risk. According to the posterior predictive checks, the model with noninformative prior distributions appeared to be somewhat better. However, the model with informative priors was also sufficiently good.

The residuals of the Hamilton model appeared to be normally or nearly normally distributed. In fact, our data set had one positive outlier which caused the rejection of a normality test. This is due to the fact that the model does not include a regime for strong boom periods of the economy. However, it is not necessary to make the model more complicated by introducing a third regime, since positive outliers are extremely rare and it is sufficient for our purpose to model the depression periods.

The fit of a model can be checked by producing replicated data sets using posterior predictive simulation. A replicated data set is produced by first generating the unknown parameters (and in the case of the Hamilton model also the states) from their posterior distribution and then, given these parameters, the new data values. One can simulate distributions of arbitrary test statistics under the checked model by calculating the test statistics from each replicated data set. Then one can compare these distributions with the statistics of the original data set. This approach to model checking is well explained in Gelman et al. (2004), in Chapter 6.

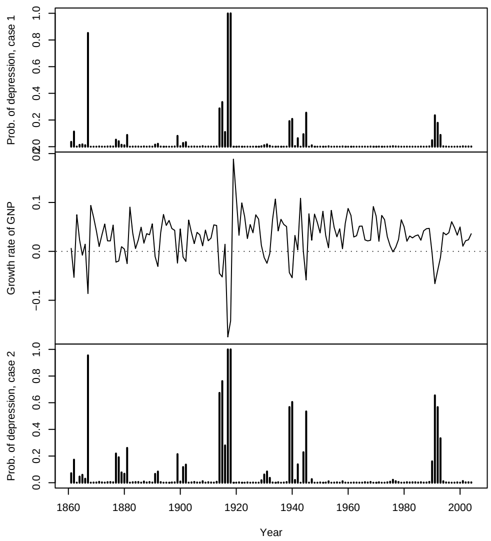

Fig. 4. The growth rate of GNP and the probabilities of depression, estimated using two different kinds of prior information. The uppermost part of the figure corresponds to the case when noninformative prior distributions are used for all parameters and the undermost part the case when informative prior distributions are used for p and q.

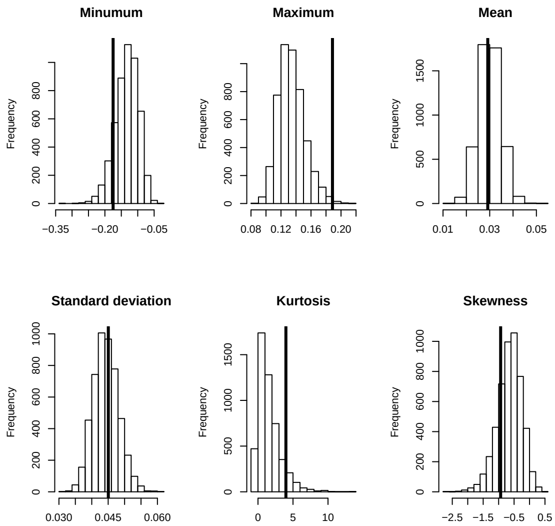

Fig. 5. Replication check for the Hamilton model with noninformative prior distributions.

We generated 5000 replicates of the GNP data under the Hamilton model and calculated some basic statistics from them. The resulting distributions were consistent with the observed statistics, which can be seen from Figure 5. Only the maximum value of the original data set is extreme with respect to its simulated posterior distribution. Still, this value is plausible under the simulated model, that is, the Hamilton model. We also made a similar test for

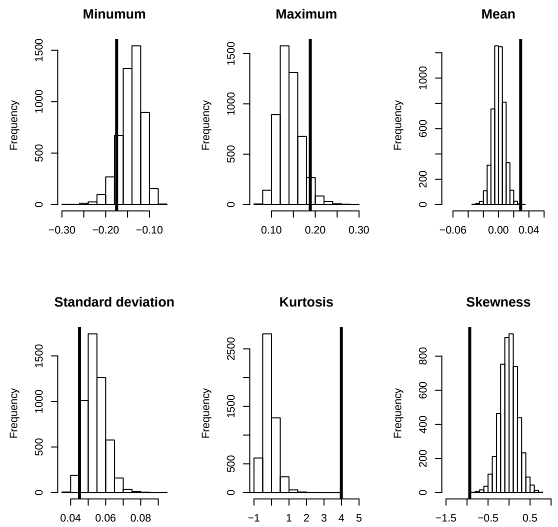

Fig. 6. Replication check for the AR(2) model.

the simpler linear AR(2) model by producing 5000 replicates. The resulting distributions, as seen in Figure 6, were not as consistent with the observed statistics as they were in the case of the Hamilton model. Specifically, the observed mean, skewness and kurtosis were more extreme than we would expect under a good model.

The discrepancy between the data and the model may be measured using several criteria; see Gelman et al. (2004). We used the average discrepancy, defined as $D_{\text{avg}}(y) = \text{E}(D(y,\theta)|y)$, the posterior mean of the deviance $D(y,\theta) = -2\log p(y|\theta)$. A smaller value of this criterion indicates a better model fit. The average discrepancy is estimated as $\hat{D}_{\text{avg}}(y) = \sum_{l=1}^{L} D(y,\theta_l)/L$, where the vectors θ_l are posterior simulations. The estimated average discrepancy for the Hamilton model with our noinformative prior distribution was $\hat{D}_{\text{avg}}(y) = -539.02$ and with our informative prior distribution $\hat{D}_{\text{avg}}(y) = -549.56$. The criterion value for the AR(2) model was $\hat{D}_{\text{avg}}(y) = -274.19$, indicating that its model fit was considerably inferior to that of the Hamilton model.

We also made robustness checks by using subsample data in estimation. The results did not change considerably, when only the first half of the data set (years 1861-1932) was used. When the second half (years 1933-2004) was used the difference between the regimes became smaller and the probability of the depression regime increased. This is natural, since the second half does not contain the years when GNP had extreme drops, that is, the years 1867 (one of the great hunger years in Finland) and 1917-18 (when Finland became independent and had the civil war).

We also made checks for our transfer function model, used in the estimation of the risk premium and the initial risk reserve. The predictive distributions of

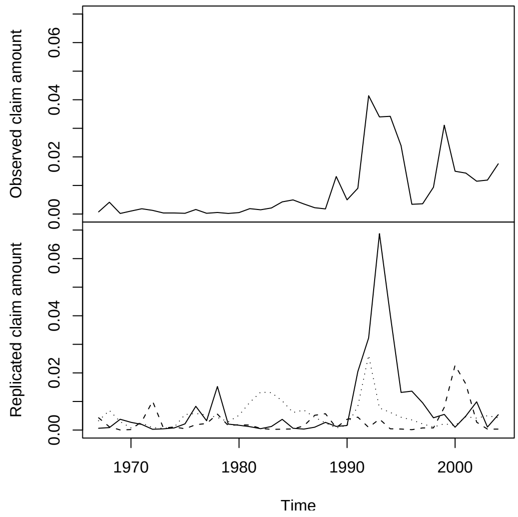

Fig. 7. The observed proportions of gross claim amount to technical provision and three replicated series when the probit transformation is used.

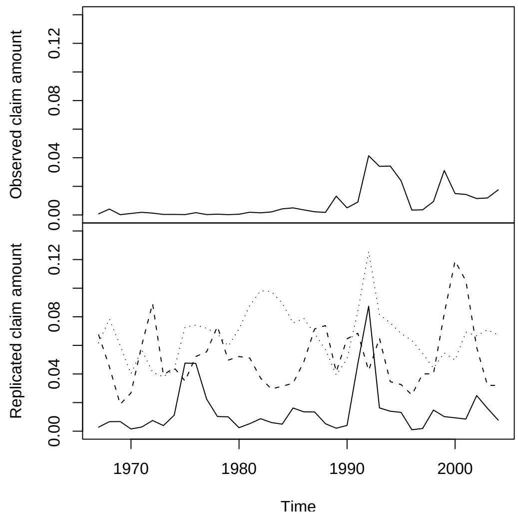

Fig. 8. The observed proportions of gross claim amount to technical provision and three replicated series when the logit transformation is used.

the basic statistics were consistent with their observed values, both in the cases of probit and logit links. The residuals, obtained after fitting the probit or logit transformed data sets, appeared to be normally or nearly normally distributed. The observed proportions of claim amount to technical provision are shown in the upper parts of Figures 7 and 8 and three replicated series in their lower parts. The probit and logit transformations are used in Figures 7 and 8, respectively, for the sake of comparison. On the basis of visual inspection the probit link would seem to be better than the logit link in producing replicated series which resemble the original series. However, one cannot draw reliable conclusions about this, since the estimation period is so short. If needed, one could take into account the model risk related to the link function by applying a more general link function, such as the inverse of the distribution function of Student's t-distribution.

A standard approach would be using a compound Poisson process to model the numbers of claims and their sizes simultaneously. However, we found such an approach difficult, since both the claim size distribution and the intensity of claims turned out to be variable during our short estimation period.

6 Conclusions

In this paper we presented an application of Bayesian modelling to financial guarantee insurance. Our goal was to model the claim process and to predict the premium and the required amount of risk capital needed for the claim deviation. Even though the data used is from the Finnish economy and from the financial guarantee system of the Finnish statutory pension scheme, we think that the model could be used in similar cases elsewhere. However, for the interpretation of the results it is important to note that the risks are probably smaller in conventional companies that operate solely on a commercial basis than in a statutory system.

The Markov regime-switching model was used to predict the number and length of depressions in the future. We used real GNP data to measure the economic growth. The claim amounts were predicted by using a transfer function model where the predicted growth rate of real GNP was an explanatory variable. We had no remarkable convergence problems when simulating the joint posterior distribution of the parameters even though the prior distributions were noninformative or only mildly informative. The sensitivity to the choice of the link function (probit or logit) in the context of the transfer function was much greater than the sensitivity to the prior assumptions (informative or noninformative) in the growth rate model.

The simulation results can be summarized as follows. First, if the effects of economic depressions are not considered properly, there is a danger that the premiums of financial guarantee insurance will be set at a too low level. Pure premium level based on the gross claim process is assessed to be at minimum 1.9%. In Finland the claim recoveries after the realization process of collaterals has been about 50%. Second, in order to get through a long-lasting depression a financial insurer should have a fairly great risk reserve. The 95% value at risk for a five-year period is about 2.3 times the five-year premium. The corresponding 75% value at risk is only about 0.30 times the five-year premium.

These figures illustrate the essential importance of reinsurance contracts in assessing the needed risk capital.

General observations can be made from this study:

- In order to understand the effects of business cycles on financial insurers' financial condition and better appreciate the risks, it is appropriate to extend the modelling horizon to cover a depression period;
- A financial guarantee insurance company may benefit from incorporating responses to credit cycle movements into its risk management policy;
- The use of Bayesian methods offers significant advantages for assessing uncertainty;
- This study underlines the observation that a niche insurance company may need special features in its internal model.

We suppose that the proposed method can also be applied to the financial guarantee and credit risks assessment of a narrow industry sector whenever a suitable business cycle model is found.

A Estimation results of the Hamilton model with noninformative prior distributions

```
Number of chains = 3
Sample size per chain = 2500

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

             Mean        SD  Naive SE Time-series SE
alpha0   0.034387 0.0037252 4.302e-05      0.0001311
alpha1  -0.127041 0.0300157 3.466e-04      0.0014282
phi1     0.272877 0.1075243 1.242e-03      0.0031602
phi2    -0.161943 0.0960893 1.110e-03      0.0027343
sigmaE   0.001325 0.0001898 2.192e-06      0.0000047
p        0.972081 0.0213633 2.467e-04      0.0010495
q        0.402894 0.2120644 2.449e-03      0.0037676
sum(St)  5.673333 3.8759548 4.476e-02      0.2744217
```

2. Quantiles for each variable:

	0.025	0.25	0.50	0.75	0.975
alpha0	0.0274775	0.031792	0.034295	0.036825	0.042004
alpha1	-0.1864689	-0.148860	-0.125884	-0.104208	-0.074866
phi1	0.0462000	0.204920	0.277852	0.345023	0.476358
phi2	-0.3557249	-0.225779	-0.158533	-0.096778	0.022228
sigmaE	0.0009913	0.001193	0.001312	0.001444	0.001734
p	0.9156113	0.962588	0.977417	0.987580	0.997095
q	0.0506376	0.236130	0.389157	0.554960	0.834136
sum(St)	2.0000000	3.000000	4.000000	7.000000	16.000000

Gelman and Rubin's diagnostics
(Potential scale reduction factors):

	Point est.	0.975 quantile
alpha0	1.01	1.03
alpha1	1.03	1.11
phi1	1.00	1.01
phi2	1.00	1.00
sigmaE	1.01	1.02
p	1.03	1.09
q	1.00	1.01
sum(St)	1.05	1.16

References

Alexander, C., 2005. The Present and Future of Financial Risk Management. Journal of Financial Econometrics, 3, 3-25.

Best, N., Spiegelhalter, D., Thomas, A., Brayne, C., 1996. Bayesian analysis of realistically complex models, Journal of the Royal Statistical Society, Series A, 159, 323-42.

Cairns, A., 2000. A Discussion of Parameter and Model Uncertainty in Insurance. Insurance: Mathematics and Economics, 27, 313-330.

Coe, P., 2002. Financial crisis and the great depression - A regime switching approach. Journal of Money, Credit and Banking, 34, 76-93.

Carlin, B. P., Polson, N. G., Stoffer, D. S. (1992). A Monte Carlo approach to nonnormal and nonlinear state-space modeling. Journal of the American Statistical Association 87, 493-500.

Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling, Journal of the American Statistical Association, 85, 972-985.

Gelman, A., Carlin, J. B., Stern, H. S., Rubin, D. B., 2004. Bayesian Data Analysis. Chapman & Hall/CRC, Second ed.

Geman, S., and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6, 721-741.

Hamilton, J., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica, 57, 357-384.

Hardy, M., 2002. Bayesian risk management for equity-linked insurance. Scandinavian Actuarial Journal, 3, 185-211.

Koskinen, L., Pukkila, T., 2002. Risk caused by the catastrofic downturns of the national economy. 27th Transactions of the International Congress of Actuaries, Cancun, Mexico.

McNeil, A., Frey, R., Embrechts, P., 2005. Quantitative risk management. Princeton Press.

Norberg, R., 2006. The Pension Crisis: its causes, possibile remedies, and the role of the regulator. Erfaringer og utfordringer Kredittilsynet 1986 – 2006, Kredittilsynet.

Pension Insurance Data Book 2005.
http://www.pbgc.gov/docs/2005databook.pdf

Rantala, J., Hietikko, H., 1988. An application of time series methods to financial guarantee insurance. European Journal of Operational Research, 37, 398-408.

Romppainen, Y., 1996. Credit insurance and the up and down turns of the national economy. XXVII Astin Colloquium, Copenhagen, Denmark.

Sigma 2000. Credit insurance and surety; solidifying commitments, 6, Swiss Re.

Smith, R., Goodman D., 2000. Bayesian risk analysis. In: P. Embrechts (eds.). Extremes and integrated risk management, Risk Books.

Scollnik, D., 2001. Actuarial modelling with MCMC and Bugs. North American Actuarial Journal, 5, 96-124.