Predictions based on certain uncertainties - a Bayesian credit portfolio approach

Christoff Gössl,
HypoVereinsbank AG, London, 14th July 2005
christoff.goessl@HVBEurope.com

Summary

The analysis of default probabilities and correlations within credit risky portfolios is usually strongly affected by the scarce availability of data. High standard deviations and a fair amount of uncertainty in the derived estimates are well known consequences of this. However, when deriving predictions in a second stage these volatilities are usually ignored and only point estimators are used, giving a false appearance of accuracy. The aim of this paper is to show how a consideration of these uncertainties will affect this second stage analysis. Besides the introduction of a new Bayesian credit portfolio approach, for this purpose in a Bayesian framework the joint posterior distribution of default probabilities and correlation parameters will be derived. Further, the effects are quantified a consideration of this distribution would have, with respect to the prediction of portfolio risk figures and also for pricing of structured derivatives.

Keywords: Credit risk; Bayes; MCMC.

Remark: The opinions expressed here are those of the author and do not reflect the views of HypoVereinsbank or its staff. Written correspondence can be sent to christoff.goessl@HVBEurope.com, or to 41 Moorgate, London EC2R 6PP.

1. Introduction

Historic default time series are a natural data source for predicting the future evolution of credit risky instruments. Besides the estimation of individual default rates, thereby especially for the assessment of portfolio risk, the estimation of potential dependencies between elements of the portfolio plays an important role. In recent publications several approaches have been presented to quantify default probabilities of different rating classes and correlations within and between industries (see e.g., Demey et al. 2004, De Servigny and Renault 2003, and Gordy and Heitfield 2003).

Two, amongst other areas in which the findings of the above approaches are utilised, are credit portfolio management and the structured finance sector. In both areas an accurate description of the risks inherent to credit portfolios is essential. There, the calculated maximum likelihood estimators for default probabilities and correlations are used to assess the future risk of a portfolio of interest, where risk is in many cases defined as a certain quantile of the portfolio loss distribution. In the latter applications, even though in most publications standard deviations or errors of estimators due to the limited sample size and data quality are well recognisably displayed, these uncertainties about the accuracy of the used parameters are usually ignored. Because of the supposedly complex distributions of and dependency structures between the parameters, in classic statistics a consideration of these relationships is also a far from straightforward task. Furthermore, this measurement error problem is the subject of a complete research area.

Bayesian models in general have already been introduced to a wide range of applications. Due to its underlying idea and theoretical simplicity, hierarchical Bayesian modelling allows for the formulation and solution of a wide range of complex questions and problems. The quantification of highly multi-dimensional distributions, stochastic restrictions, as well as the incorporation of temporal and spatial dependency structures build the foundation of the recent success story of the Bayesian paradigm. Moreover, the benefits of the incorporation of prior information to mitigate the consequences of sparse data bases and measurement error environments are a well known fact. In this paper the main focus lies on the development of a credit portfolio model within a Bayesian framework. One advantage of such a Bayesian approach would be, that one would easily be capable of overcoming the above mentioned problems. Using adequate estimation techniques, within a Bayesian framework it is possible to derive and make available the joint distribution of a set of parameters, conditional on a given set of observations, and not only a few descriptive statistics. This distribution could then be used in a second inferential stage to consider the randomness of derived default probability and correlation estimates and to assess what the effects are with respect to risk figures and pricing.

The aim and the structure of this paper is twofold. The first part briefly recapitulates the basic principles of the Bayesian approach, with emphasis on the fundamental switch of paradigm inherent in it. Here a short overview of the standard estimation approach for Bayesian models, the Markov Chain Monte Carlo (MCMC) methods, will also be given. In the second and main part a first simple application of the Bayesian paradigm to the credit risk sector will be proposed to show its potential. After the formulation of a Bayesian standard one factor credit portfolio model, estimates for individual rating dependent probabilities of default (PDs) and the portfolio correlation are derived from rating agency data. As this could also easily be done within the classical framework, additionally the joint dependency structures between these parameters are quantified and the benefits and implications are shown these results will have when applied to the calculation and prediction of risk figures for particular portfolios.

As a last remark, throughout the paper, all densities will be denoted by $p(.)$ and all distribution functions will be denoted by $P(.)$. The particular contexts should reduce potential confusions regarding the corresponding probability spaces to a minimum.

2. The Bayesian Approach

The fundamental idea of Bayesian statistics builds on the Bayesian theorem

$$\begin{aligned} p(\beta|data) &= \frac{p(data,\beta)}{p(data)} \\ &= \frac{p(data|\beta)\; p(\beta)}{\int p(data|\beta)\; d\beta}\,, \end{aligned} \tag{1}$$

describing the relation between a set of observed data and some unknown parameters β. Inherent to the theorem is the basic difference between Bayesian and classical approaches in statistics, the assumption that the model parameters though unknown are nevertheless no longer fixed. It is supposed that these parameters have a distribution, too, comparable to the observed data sample. And this distribution can and has also to be specified. We will come back to the problems and opportunities this requirement imposes in a moment. This so called prior distribution or information is now combined with the information about the parameters contained in the data sample to form a so called posterior distribution of parameters conditional on the observed data. This is sometimes also called a learning process or an update of the prior information with the data sample. The resulting posterior distribution finally holds all information regarding the parameters, contributed once by the priors and once by the observed data. To calculate the posterior, one part is built by the standard likelihood function of the observed data conditional on the unknown parameters, the other part is covered by the prior of the parameters.

For the formulation of these priors two general cases can be distinguished. The first is where substantial prior information is available, i.e. specific distributions or samples of previous analyses, and can be expressed as a distributional assumption, not necessarily in an analytical or parametrical fashion. These are evidently a natural choice for prior distributions. However, the second and more common setting is where there is only very little or no prior information available. In this case this non-existent information has also to be translated into a distributional assumption. For this purpose usually overdispersed or flat priors are applied, e.g., uniform distributions on the unit interval for probabilities $p \sim \mathbf{U}(0,1)$, or flat normal distributions for metric unrestricted parameters $a \sim \mathbf{N}(0, s^2)$ with a sufficiently large variance s^2 ($\gg$1000) to allow for a practically overdispersed distribution on a reasonable interval for a. To allow in not too complex settings for an analytic solution for the posterior, also so called conjugate priors can be used. For normally distributed observations, for example, normal and inverse gamma prior distributions for the unknown mean and variance result in the same posteriors for the respective parameters. But this should be mentioned for completeness only, due to the introduction of computerintensive numerical estimation techniques one is more or less free in the choice of the priors to apply.

As a general remark, it is also the formulation of the prior distributions that accounts most towards the flexibility of the Bayesian approach. Smoothly over time changing parameters can be described by autoregressive random walk priors. Two dimensional versions of these can also be used to introduce spatial dependencies into the estimation, where differences in the parameter values of neighboring observations are penalised, yielding a smooth parameter surface. Robust spatial priors, using Cauchy or Student distributions, can even be used to describe edges within the surface. And as a matter of fact, the use of latent variables with adequate prior distributions to achieve conditional independence, as also applied in the factor models for credit portfolios, is one of the most frequently used ideas in Bayesian statistics.

3. Estimation – MCMC methods

Even though the Bayesian theorem has been known for over 200 years and also the Bayesian paradigm of a learning process and its advantages have been lively discussed in statistics, a major drawback until the 1980s–1990s was the fact that apart from some simple and analytically tractable settings, the denominator of theorem (1) can not be determined. The posterior can be quantified only up to a normalising constant, making it almost impossible to analyse its characteristics because standard direct Monte Carlo methods like rejection or importance sampling can not be applied. Even though the principles of algorithms which were able to overcome this limitation were already presented by Metropolis et al. (1953) and Hastings (1970), only the dramatic increase in computational power finally allowed the application in statistics. Once started by publications of e.g. Besag et al. (1991), Smith and Roberts (1993), or Gilks et al. (1996), the MCMC techniques soon became the method of choice in Bayesian statistics because of their flexibility, robustness, and almost unlimited applicability.
In contrast to direct Monte Carlo methods, MCMC techniques do not sample from a distribution or density of interest directly but construct a stationary Markov Chain whose transition kernel converges against this distribution. Once convergence has been reached, realisations of the chain are at the same time realisations of this (posterior) distribution. Thereby, it is fully sufficient if the distribution in question can only be specified up to a normalising constant. From this sample all characteristics of interest with respect to the posterior can be derived. A theoretically more profound introduction and practical applications can be found in Casella and George (1992), Tierney (1994), and Chib and Greenberg (1995). Further, an important aspect worth mentioning is that the realisations represent a sample from the joint posterior distribution of the parameters, allowing also for conclusions regarding higher order dependencies between parameters.
Many Bayesian models can usually be implemented and estimated using freely available software packages, as e.g. WinBUGS (Spiegelhalter et al., 2003) or BayesX (Lang et al., 2004). Within these packages samples from the posterior distributions are generated using Gibbs or Metropolis Hastings algorithms. Further, advanced monitoring tools are supplied to ensure mixing properties and convergence of the used Markov chains. Sample means or medians are well established point estimators for the unknown parameters. If overdispersed priors are applied, taking the posterior mode as estimator would yield comparable results as in the maximum likelihood framework. However, because in practice even for massive samples in high dimensional settings the mode is empirically quite

difficult to determine exactly, usually the above statistics are used. Highest posterior density regions can be used as the Bayesian analogue to classical confidence intervals.

4. A simple Bayesian credit portfolio model

In the second part of this paper the Bayesian idea will be applied to the credit risk sector. In the following let us consider a portfolio of N credit risky instruments or assets $a_i,\ i = 1, \dots, N$, each comprising one unit. Each of these assets can be classified into one of a finite number of risk classes $r_i \in \{R_1, \dots, R_K\}$, defining its risk profile completely. The risk or rating classes have individual one year probabilities of default $p_{R_j},\ j = 1, \dots, K$, with the asset PDs $p_i = p_{R_j}$ for $r_i = R_j$. Further, let us suppose the respective instruments follow a one factor asset value model, as introduced by Merton (1974) or Vasicek (1991). In this framework it is assumed that an instrument or firm defaults when its asset value process X_i falls below the firm's liabilities or a certain default frontier. The corresponding default frontier or threshold is determined by its individual risk profile or risk class r_i. In its discrete version on a one year horizon, the underlying asset return process is defined as follows:

$$\begin{aligned} X_i &= \sqrt{\rho}\, Y + \sqrt{1-\rho}\, Z_i\ , \quad i = 1, \dots, N\ , \\ & \rho \in [0,1],\ Y \sim \mathbf{N}(0,1),\ Z_i \sim \mathbf{N}(0,1)\ \ i.i.d.\ . \end{aligned} \tag{2}$$

The process is the sum of a common factor Y invariant throughout the portfolio and an idiosyncratic component Z_i resembling an independent individual contribution of asset i to its evolution over time. These two components are weighted by a correlation parameter ρ determining the intra-portfolio dependencies within the whole portfolio. This model also corresponds to Gordy and Heitfield's restrictions R1 and R3. Even though there might be criticisms about the adequacy of this simplifying approach, however, for the ideas to be shown in this article it is perfectly well suited. Potential generalisations to more flexible multi factor approaches should be obvious. With the above default probabilities the following relation can be written down,

$$\begin{aligned} P(\text{ asset i defaults }) &= P(X_i < k_i) \;=\; p_i \\ \Longrightarrow \quad k_i &= \Phi^{-1}(p_i)\ . \end{aligned} \tag{3}$$

Further, conditional on the portfolio factor Y this equation becomes

$$\begin{aligned} P(\text{ asset i defaults } \mid Y = y\ ) &= P(X_i < k_i \mid Y = y) \;=\; p_{i|y} \\ \Longrightarrow \quad p_{i|y} &= \Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\rho} \times y}{\sqrt{1-\rho}}\right)\ , \end{aligned} \tag{4}$$

completing a generalised linear model for Bernoulli data with a probit link function. Utilising the conditional independence of the elements on the factor Y and the grouping into homogeneous risk classes allows further to specify the joint probability distribution of observing a certain number l_{R_j} of defaults in the respective risk classes of the portfolio comprising n_{R_j} assets,

$$\begin{aligned}
L_{j|Y=y} &= \sum_{r_i=R_j} \mathbf{1}_{\mathbf{D_i}|\mathbf{Y=y}}, \quad j=1,\dots,K, \\
P(L_1=l_1,\dots,L_K=l_K|\ Y=y\ ) &= \prod_{j=1}^{K} \mathbf{B}(n_{R_j}, p_{j|y}, l_{R_j})\ ,
\end{aligned}$$

with $\mathbf{B}(n,p,k)$ the standard Binomial distribution. Without loss of generality and for simplicity at this point it is supposed that in case of default every obligor suffers the same loss of one unit, i.e. their loss given default is 100%. General concepts of loss severity or loss given defaults should not be too difficult to incorporate. Defining $\mathbf{L_t} = (L_{t,1},\dots,L_{t,K})$, $\mathbf{n_t} = (n_{t,1},\dots,n_{t,K})$, $\mathbf{p} = (p_{R_1},\dots,p_{R_K})'$, and also considering the other unknown parameters the complete conditional likelihood function for period t and a one year horizon can be rewritten in a concise form

$$P(\mathbf{L_t} = \mathbf{l_t} \mid \mathbf{n_t}, \mathbf{p}, \rho, y_t\ ) = \prod_{j=1}^{K} \mathbf{B}(n_{t,R_j}, p_{j|y_t}, l_{t,R_j}). \tag{5}$$

As pointed out in section 2, in addition to the likelihood function, to complete the Bayesian portfolio model it is left only to specify the prior distributions. Assuming the case of no substantial prior information and therefore flat priors for the parameters, and exploiting the serial independence of observations, this yields the complete hierarchical Bayesian credit portfolio approach and the corresponding posterior distribution of the unknown parameters conditional on the observations $(\mathbf{n}, \mathbf{l})$

$$\begin{aligned}
P(\mathbf{p}, \rho, \mathbf{y} \mid \mathbf{n}, \mathbf{l}) &= \frac{\prod_{t=1}^{T} P(\mathbf{L_t} = \mathbf{l_t} \mid \mathbf{n_t}, \mathbf{p_{y_t}}, \rho, y_t\ )\ \ p(\mathbf{p})p(\rho)p(\mathbf{y})}{P(\mathbf{L} = \mathbf{l} \mid \mathbf{n})}\ , \qquad (6) \\
\mathbf{p_{y_t}} &= (p_{1|y_t},\dots,p_{K|y_t})'\ , \\
p_{j|y} &= \Phi\left(\frac{\Phi^{-1}(p_{R_j}) - \sqrt{\rho} \times y}{\sqrt{1-\rho}}\right)\ , \\
& \quad p_{R_j} \sim \mathbf{U}(0,1),\ i.i.d.,\ j=1,\dots,K\ , \\
& \quad \rho \sim \mathbf{U}(0,1)\ , \\
& \quad y_t \sim N(0,1),\ i.i.d.,\ t=1,\dots,T\ ,
\end{aligned}$$

thereby defining $\mathbf{U}(0,1)$ as the uniform distribution in the unit interval and $\mathbf{y} = (y_1,\dots,y_T)'$, and $P(\mathbf{L} = \mathbf{l} \mid \mathbf{n})$ can be calculated by integrating the numerator of equation (6) with respect to $\mathbf{p}$, ρ, and $\mathbf{y}$. Indicated by the priors above, in this paper the correlation between instruments is assumed to be positive, even though this is known to be the subject of a lively discussion. The above model is easily estimated using the WinBUGS package and in the remainder of the paper its application to historic Standard & Poors rating agency default data will be discussed in detail.

5. Application – Analysis of historic data

The data which will be used to illustrate the Bayesian approach were published in Standard & Poors' default report (Standard & Poors, 2005). In this report for the last 24 years, beginning in 1981, the numbers of rated companies are listed and also the fraction thereof defaulting, split up by their respective rating categories. Even though agency data are also available for AAA and AA rating classes, the analysis is restricted to the lower rating grades beginning with the A class. By doing so the difficult question can be avoided, whether the lack of observing a significant number of defaults in these higher classes is the result of a very conservative classification, or whether it is the actual consequence or manifestation of correlation. It should also be pointed out that the real data examples in this article are used for illustrative purposes only and should be understood as such. No efforts were undertaken in validating the data beforehand, but the published default data were assumed to be suitable for the presented model, i.e. fulfill the assumptions made in the previous section. With these data, the initial aim is to derive the joint posterior distribution of historic average default frequencies of the rating classes, the portfolio correlation, as well as the portfolio factor time series.

We generated joint posterior distribution samples of 25,000 observations, using Markov chains of a length of 2,200,000, and taking every 80th state of the chains, with a burn in phase of 200,000 iterations. By this strategy autocorrelations and convergence diagnostics of the algorithm reached sufficiently low levels. Table 1 shows a number of descriptive statistics of the posterior distributions. Figure 2 displays two-dimensional contour plots, as well as kernel density estimates for selected two-dimensional marginal distributions. Further, in Figure 1 the posterior mean of the portfolio factor time series is shown together with its 97.5% and 2.5% quantiles.

It can easily be verified that for rating classes A to B the results for the average PDs are slightly higher than the average number of defaults over time. Correlation effects and the use of the arithmetic mean as a Bayesian point estimator should be the most likely reasons for these differences. Especially for skewed distributions such as the observed (see e.g. Figure 2c,d), shifts of the arithmetic mean towards the heavier tails are to be expected. Further, the estimate for the correlation factor is roughly at the same level compared to the results of Gordy and Heitfield for restrictions R1 and R3, and the results of De Servigny and Renault. Besides the already mentioned explanations, changes in the underlying data and the sensitivity of this parameter even to slight data changes should contribute most to the differences. As a backtest, also the approximate Bayesian posterior mode has been derived for the data used by De Servigny and Renault, with similar results compared to their estimated portfolio correlation of 6.3%. For these numbers, due to the highly skewed correlation distribution the posterior mean was at 7.9%. However, particularly noteable are the high standard deviations especially for the higher rating classes and also for the correlation parameter, indicating a fair amount of uncetrtainty in these estimates. The time series of portfolio factors themselves are not remarkably different to the results when solving equation (4) for the variable Y, using frequentistic estimates for the correlation and for example the PD estimate for class CCC. Nevertheless, at least one important difference exists between classical solutions for this problem and the Bayesian results. For all mentioned parameters a sample of their joint posterior distribution exists. Especially with respect to the portoflio factors, parameters do not have to be estimated using a back-fitting algorithm or a two step approach, they can be derived simultaneously. Instead of being confined to the Hessian or information matrix, the above sample allows further to derive a whole variety of characteristics of their joint dependence structure, such as joint excedence probabilities etc., or simply to resample from it. For the remainder or this paper, the latter aspect is of special interest. The problem of how to account for the estimators' volatilities in the prediction simplifies dramatically with their joint distribution available. It reduces to a simple resampling problem.

6. Application – Predictions

One of the most important applications of the above results is the prediction of risk levels for particular portfolios. When defining risk as deviations from or uncertainties about expectations, in addition to variations in external regressors, one should also consider the uncertainties built into the parameters these expectations are based on. In classical approaches, however, in this second stage usually point estimators derived from the above historical data are used to assess this risk, in general not accounting for any measurement problem. As pointed out, in the Bayesian framework this could be easily overcome because of the direct availability of a posterior distribution sample for the parameters. To show the remarkable effects an incorporation of the mentioned uncertainties can have, the default/loss distribution will be derived for homogeneous portfolios for all rating classes $r_i \in \{R_1, \ldots, R_K\}$, consisting of an infinite number of identical assets. To calculate the fraction of defaulting assets, in one case the historical point estimators are used and in the other case random samples from the joint posterior distribution are drawn. Thus, for the fraction l_{R_k} of defaulting assets in a portfolio of assets from risk class R_k equation 4 can be applied in the following contexts:

$$
\begin{aligned}
l_{R_k}(p^*, \rho^*, y) &= \Phi\left(\frac{\Phi^{-1}(p^*) - \sqrt{\rho^*} \times y}{\sqrt{1-\rho^*}}\right) , \\
& \quad y \sim N(0,1) \\
\text{I} \quad \text{classical approach :} & \quad p^* = \hat{p}, \;\; \rho^* = \hat{\rho} && (7) \\
\text{II} \quad \text{Bayesian approach :} & \quad (p^*, \rho^*) \sim P(p_{R_k}, \rho \mid \mathbf{n}, \mathbf{l}, \mathbf{y}) && (8)
\end{aligned}
$$

Thereby, to result in identical expected losses for the portfolios as point estimators $\hat{p}$ and $\hat{\rho}$, the arithmetic means are calculated from the posterior distributions, from which also the samples are drawn. Samples of 100,000 realisations have been generated. Tables 2 and 3 show some descriptive statistics of the default distributions, Figure 3 displays the corresponding kernel density estimates. In Figure 4 the differences in the quantiles, also observable in the above tables, are illustrated by normalizing them to the quantiles for approach I. In the tables and in the quantile plot considerable increases of up to 20% in the 95% and 99% quantile values can be observed, moving from the classical to the Bayesian approach. This observation should be a direct consequence of the variation of the underlying parameters in the Bayesian approach and of the dependence between PDs and correlation parameter ρ as it can also be seen in Figure 2.

From the above results for the credit portfolio modelling two conclusions can be drawn. The first is, incorporating the uncertainties generally inherent to point estimators into a portfolio model in a second stage yields significantly increased quantiles and thus risk figures in the predictions. And as a second result, this effect varies with the amount or lack of information available. In other words, the more reliable the point estimators are the less additional risk is introduced in the predictions when accounting for their randomness. The effect is largest for the A rating class and decreases when going down the rating scale. Where there are only few defaults observable for a rating class the estimated PDs standard deviation or uncertainty is quite a bit higher than for rating classes with very frequent defaults like the CCCs. Accordingly, the risk or unexpected number of defaults increases, too. It should further be mentioned that we used the posterior mode

as a point estimate in the classical approach to derive the above numbers. Due to the skewed posterior distributions, with the application of the posterior mode the increases in the quantiles should be even more prominent.
Are the above results interesting from a portfolio manager perspective, what are the effects with respect to structured products, which also heavily depend on accurate portfolio modelling? Let us consider an arguably hypothetical transaction based on one of the above infinitely granular portfolios, e.g. comprising BB assets. For simplicity and to avoid any complexities regarding multi-year PDs, we want to sell protection for one year on a tranche with an attachment point of 1.90% (i.e. the 50bps quantile of the classic prediction!) and a 1% thickness. Further, we are only concerned about the credit risk in this transaction. What would be the adequate price, i.e. the expected loss of this transaction. Per definition in the standard setting we have a hitting probability of 50bps and the expected loss can be quantified with 23bps, resulting in a loss given default of 46%. Deriving these quantities for the Bayesian approach, the PD of this tranche turns out to be at 88bps, the expected loss increases to 50bps, thus yielding a loss given default of 57%. In other words, the incorporation of some 'doubts' about the accuracy of the underlying parameters more than doubles the adequate price at which this protection should be sold. The effects on the rating of this tranche do not have to be mentioned.

7. Discussion

The aim of this article was to introduce a Bayesian approach to the modelling of credit risky portfolios. A Bayesian portfolio model was formulated, which allows to describe default frequencies and intra-portfolio correlations. Further the implications could be shown of considering the randomness of derived parameter estimates in a second stage when assessing portfolio risk, from a portfolio management point of view as well as for the purpose of pricing and rating credit portfolio derivatives.

Especially in the area of pricing and rating credit derivatives in recent years publicly available models like the S&P CDO Evaluator or Fitch's VECTOR model became more and more popular, and can without any question be regarded as an industry standard commonly accepted. However, they exactly stand for the classic approach discussed in the previous section 6. As well as for ratings as for correlations, undoubtedly with more sophisticated underlying factor models, fixed estimates or parameters are used without any consideration of their randomness, not to mention their dependence structure be-

tween each other. Additionally, the fact that many predictions are not only made for the next year, as in the above example, but usually for time periods of 5 to 7 years makes the situation even more precarious. With an increasing number of parameters to estimate or calibrate, the effects of small sample sizes usually present in this area should become increasingly important. This holds for multi-year PDs as well as for say industry specific correlation levels. The latter might possibly be derived from the broader data basis of equity data, nevertheless also due to the transfer from equity to asset or default correlations, this should not solve the problem completely.
When looking at the results of the portfolio model in table 1, as pointed out an interesting detail is the relation between the sample means and the corresponding standard deviations. For the higher rating classes and the correlation parameter the standard error amounts of 30% to 50% of the mean. Comparable results can also be seen in De Servigny and Renault (2003). The results are also in agreement with a general study on confidence intervals for default probabilities of Hanson and Schuermann (2005). An effect of this relation surfaces when omitting one or two years in the analysis. In general it is assumed that the default data are serielly independent, so this should pose no problem. However, changes of 5% to 10% especially in the correlation parameter are not unusual. This raises again the question whether the use of relatively volatile point estimators are a sensible choice to derive quantities down to precisions in the basispoint area. In Bayesian approaches as well as classical measurement error models these problems surely do not vanish, but at least it is accounted for.

Models with latent variables are one of a variety of Bayesian applications in statistics. In econometrics autoregressive hierarchical Bayesian models have already been applied to macroeconomic problems such as unemployment durations or flat rent data. Further, the use of substantial prior informations for credit scoring has been analysed. As presented in the previous sections, also in the credit risk sector Bayesian techniques proved to be useful. Perspectives for the future may therefore be a more thorough investigation, how more of their advantages can be exploited in particular for the credit risk area. For example the extension towards a multi factor approach is obviously at hand. Further, the use of time series techniques to describe the temporal evolution of underlying portfolio factors are conceivable. Moving from a copula induced single period factor model to temporal dependencies seems to be a reasonable idea. And also the incorporation of a time-varying seasonal component into the probabilities of default to put more emphasis on their point in time component could be a project for the near future.

Acknowledgements

The author wants to express special thanks to Mathieu Guinnepain for extensive and invaluable discussions on the subject of correlations and to Christopher Thorpe and Christian Bluhm for their fruitful comments.

References

Besag, J., York, J. and Mollie, A. (1991). Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* **43, 1**, 1–59.

Casella, G. and George, E. (1992). Explaining the gibbs sampler. *The American Statistician* **46**, 167–174.

Chib, S. and Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The American Statistician* **49**, 327–335.

Demey, P., Jouanin, J.-F., Roget, C. and Roncalli, T. (2004). Maximum likelihood estimate of default correlations. *Risk,* **November 2004**, 104–108.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.

Gordy, M. and Heitfield, E. (2003). Estimationg default correlations from short panels of credit rating pperformance data. *Federal Reserve Board* .

Hanson, S. and Schuermann, T. (2005). Confidence intervals for probabilities of default. Technical report, Federal Reserve Bank of New York.

Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika* **57**, 97–109.

Lang, S., Brezger, A. and Kneib, T. (2004). *BayesX – Software for Bayesian Inference*. Department of Statistics, University of Munich, http://www.stat.uni-muenchen.de/ bayesx.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance* **29**, 449–470.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics* **21**, 1087–1091.

Servigny, A. D. and Renault, O. (2003). Correlation evidence. *Risk,* **July 2003**, 90–94.

Smith, A. F. M. and Roberts, G. O. (1993). Bayesian computation via the gibbs sampler and related markov chain monte carlo methods (with discussion). *Journal of the Royal Statistical Society, Series B* **55**, 3–23.

Spiegelhalter, D. J., Thomas, A., Best, N. and Lunn, D. (2003). *BUGS Software*. MRC Biostatistics Unit, Institute of Public Health, Cambridge, http://www.mrc-bsu.cam.ac/bugs.

Standard & Poors (2005). Annual global corporate default study: Corporate defaults poised to rise in 2005. Technical report, Global Fixed Income Research.

Tierney, L. (1994). Markov chains for exploring posterior distributions. *The Annals of Statistics* **22, 4**, 1701–1762.

Vasicek, O. A. (1991). The loan loss distribution. Technical report, KMV Corporation.

	A	BBB	BB	B	CCC	rho
Hist. Mean	0,0004	0,0023	0,0107	0,0561	0,2813	
Mean	0.0006	0.0029	0.0123	0.0591	0.2789	0.0941
Std. Dev.	0.0003	0.0010	0.0031	0.0095	0.0257	0.0357
2.5%	0.0002	0.0016	0.0079	0.0445	0.2328	0.0449
25%	0.0004	0.0022	0.0102	0.0526	0.2613	0.0690
Median	0.0005	0.0027	0.0118	0.0578	0.2772	0.0873
75%	0.0007	0.0034	0.0137	0.0642	0.2949	0.1115
97.5%	0.0014	0.0055	0.0198	0.0817	0.3341	0.1826

Table 1: Descriptive statistics for the parameter posterior distributions of the Bayesian Credit Portfolio Model (6).

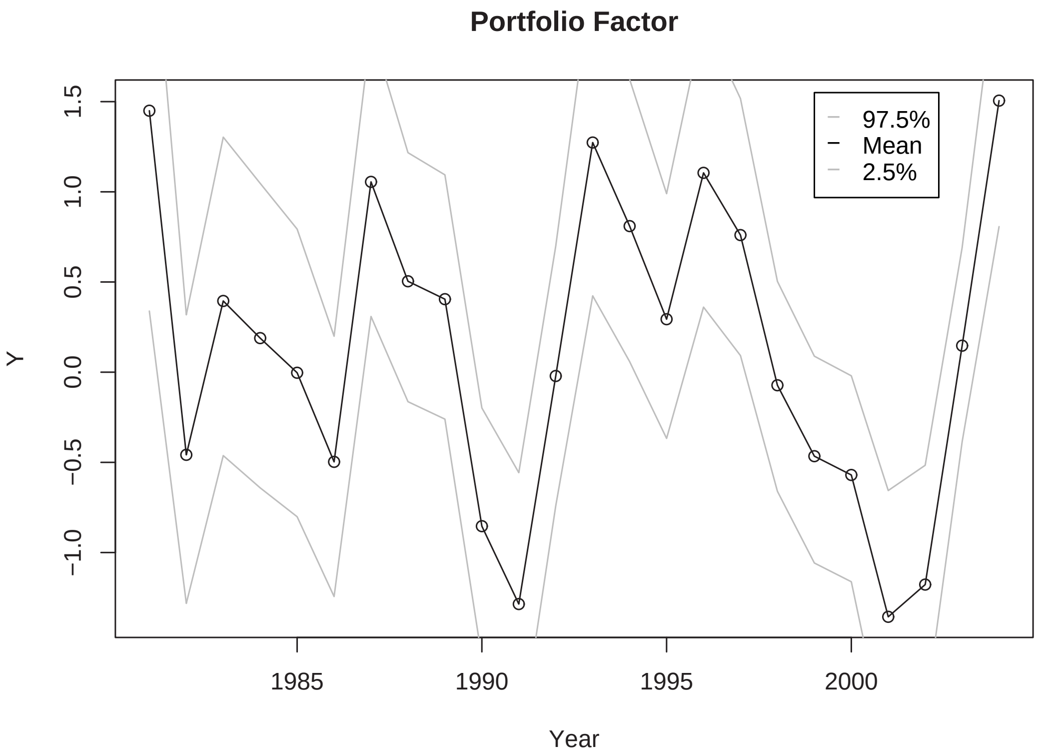

Figure 1. Posterior mean and quantiles for the portfolio Factor Y.

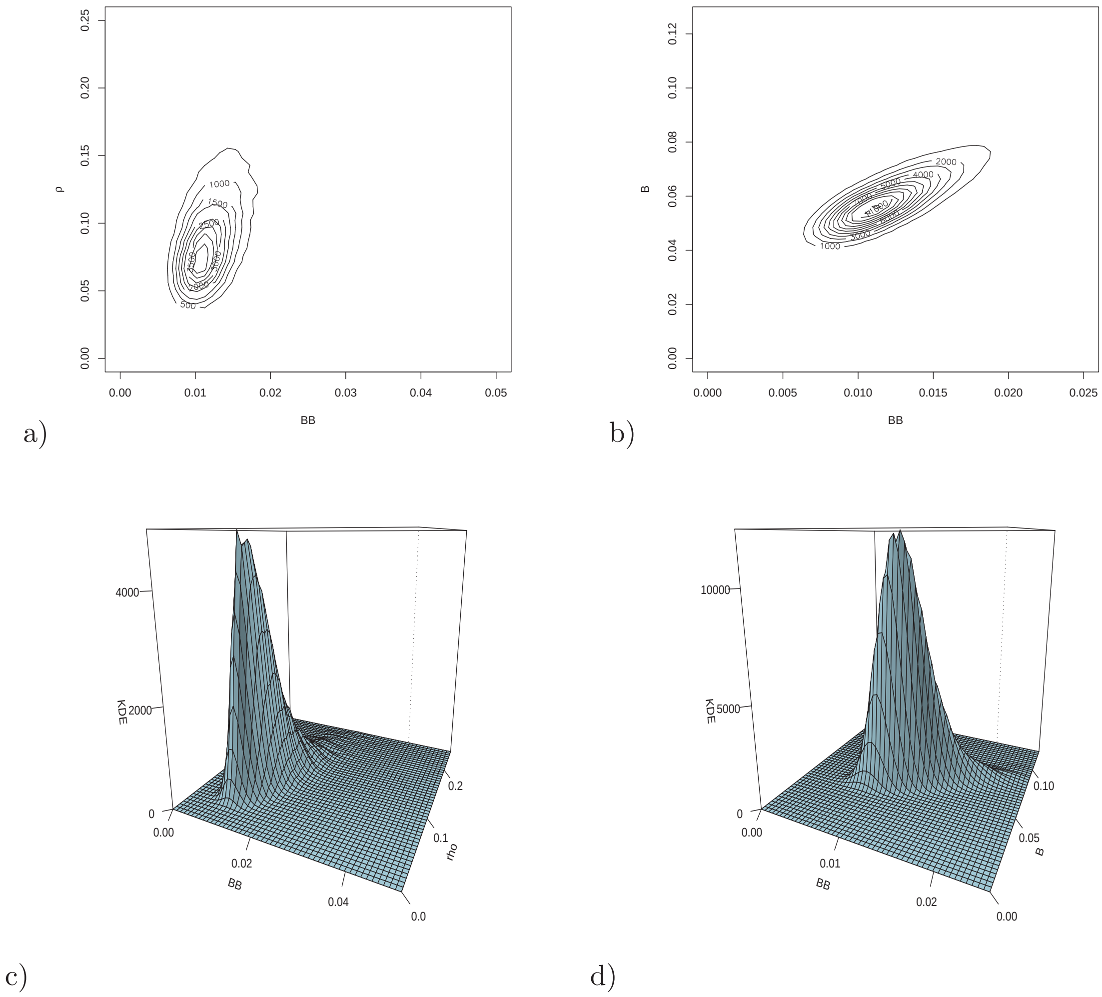

Figure 2. Joint posterior distribution for selected parameters of the Bayesian credit portfolio model (6).

	A	BBB	BB	B	CCC
Mean	0.0007	0.0041	0.0143	0.0613	0.2929
Std. Dev.	0.0008	0.0039	0.0112	0.0354	0.0969
1%	0.0000	0.0003	0.0016	0.0115	0.1067
5%	0.0001	0.0006	0.0029	0.0187	0.1481
10%	0.0001	0.0009	0.0040	0.0237	0.1735
25%	0.0002	0.0016	0.0066	0.0355	0.2223
Median	0.0004	0.0029	0.0113	0.0536	0.2851
75%	0.0009	0.0052	0.0187	0.0791	0.3556
90%	0.0015	0.0086	0.0284	0.1086	0.4238
95%	0.0021	0.0114	0.0363	0.1291	0.4643
99%	0.0039	0.0184	0.0547	0.1749	0.5408

Table 2: Predictions of risk class dependent loss distributions for infinitely granular homogenous portfolios using the classical approach (7).

	A	BBB	BB	B	CCC
Mean	0.0007	0.0041	0.0143	0.0613	0.2929
Std. Dev.	0.0010	0.0045	0.0123	0.0377	0.0998
1%	0.0000	0.0002	0.0013	0.0100	0.0958
5%	0.0001	0.0006	0.0028	0.0180	0.1445
10%	0.0001	0.0008	0.0039	0.0236	0.1728
25%	0.0002	0.0015	0.0065	0.0354	0.2229
Median	0.0004	0.0028	0.0110	0.0531	0.2844
75%	0.0008	0.0051	0.0182	0.0778	0.3533
90%	0.0015	0.0083	0.0281	0.1079	0.4239
95%	0.0022	0.0116	0.0365	0.1320	0.4719
99%	0.0046	0.0213	0.0609	0.1907	0.5682

Table 3: Predictions of risk class dependent loss distributions for infinitely granular homogenous portfolios using the Bayesian approach (8).

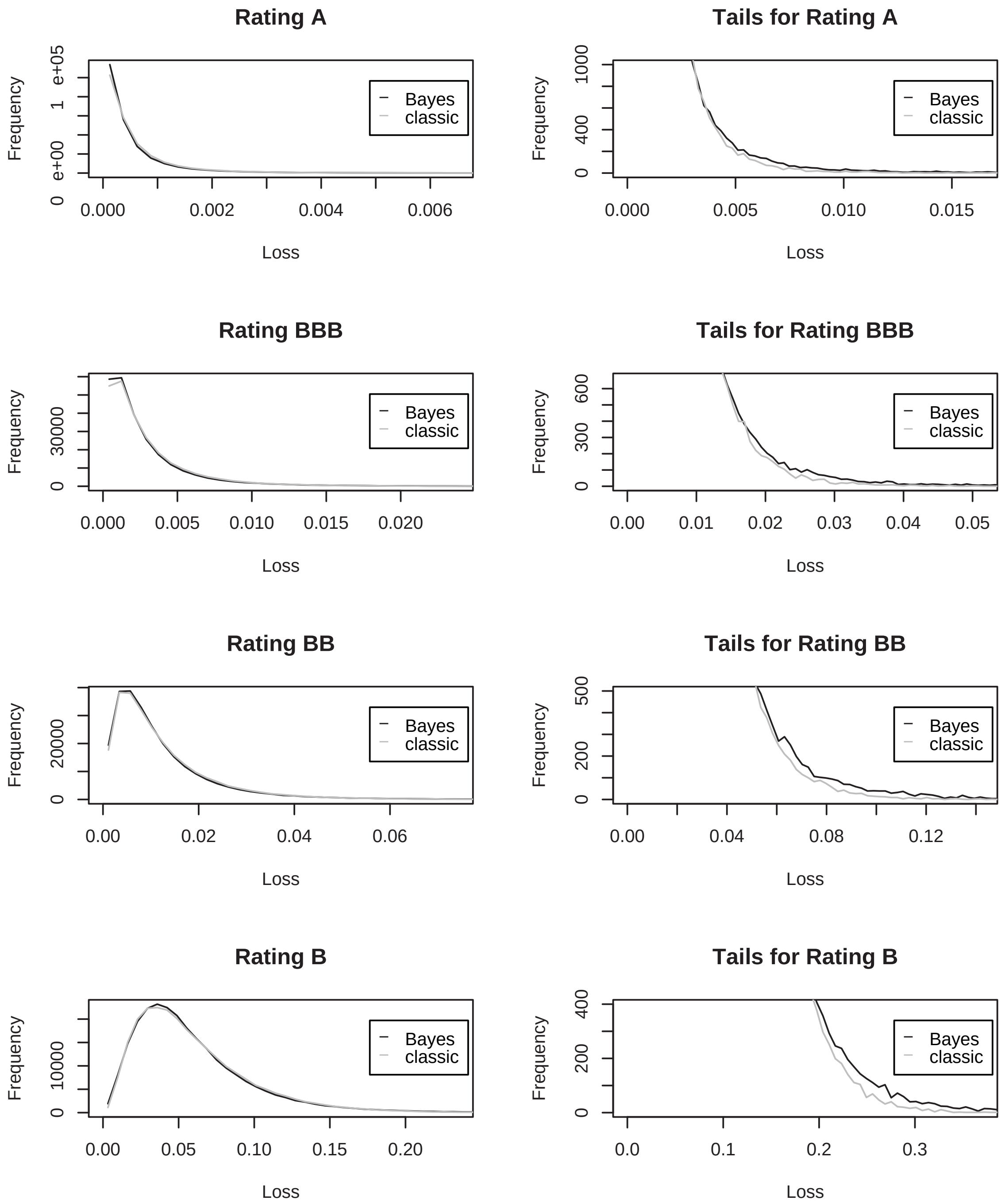

Figure 3. Density estimates for risk class dependent portfolio loss predictions. Left: full distributions. Right: zoomed tail section.

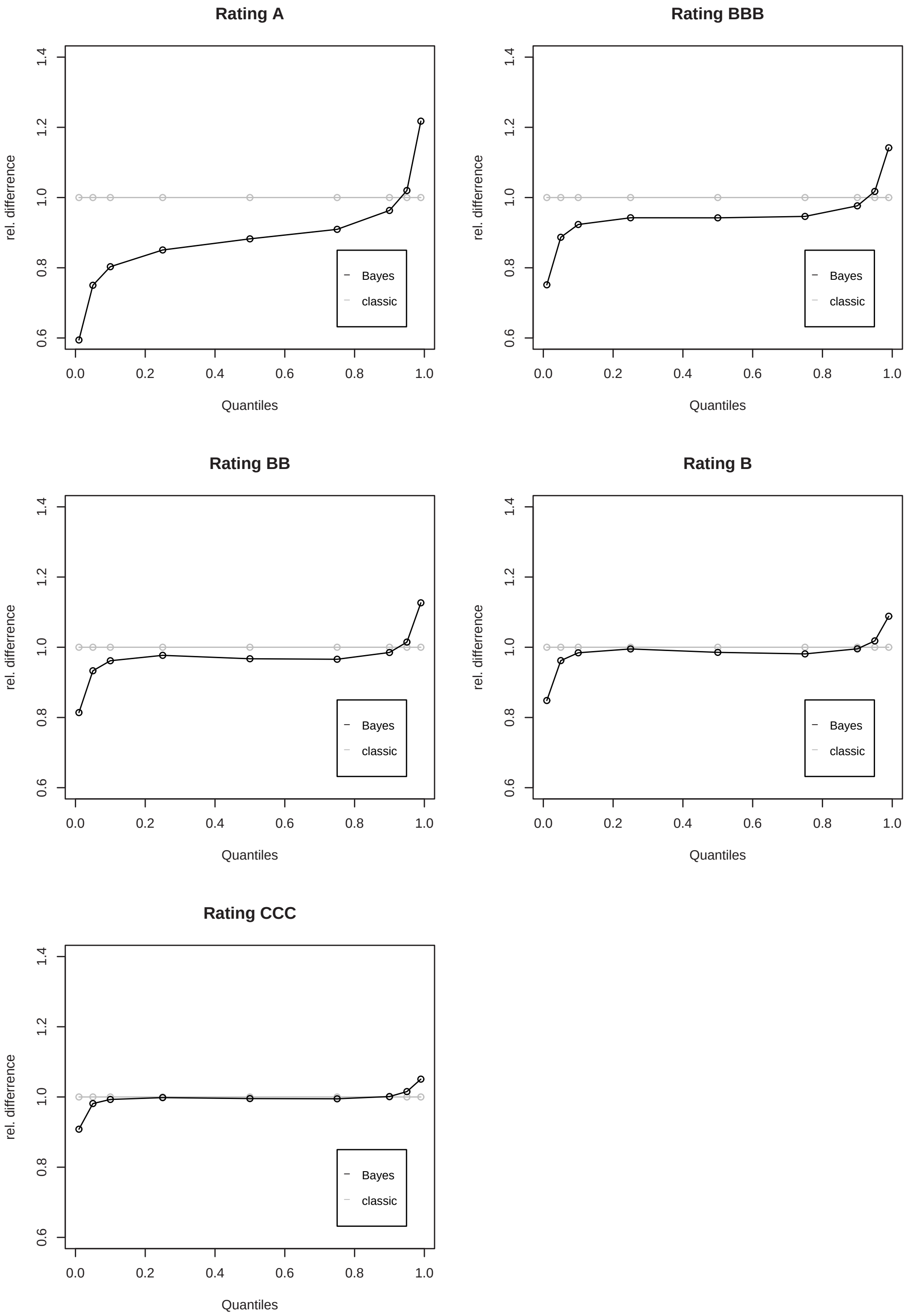

Figure 4. Comparison of quantiles for risk class dependent portfolio loss predictions, normalised to the results for the classical approach.