Forthcoming, The Journal of Banking and Finance

THE EFFECTS OF ESTIMATION ERROR ON MEASURES OF PORTFOLIO CREDIT RISK

Gunter Löffler*
University of Frankfurt (Main)

This version: October 15, 2001

Abstract

This paper uses Monte Carlo simulations to assess the impact of noisy input parameters on the accuracy of estimated portfolio credit risk. Assumptions about input quality are derived from the distribution of historical sample statistics commonly used in default risk modelling. The resulting estimation error in the distribution of portfolio losses is considerable. Losses that are judged to occur with a probability of 0.3% may actually occur with a probability of 1%. The paper also shows that estimation error leads to biases in VaR estimates and significance levels of backtests. The biases can be corrected by analysing predictive distributions which average over the unknown parameter values.

JEL classification: G21, C13

Key words: credit risk, estimation error, value at risk, predictive distributions.

* Gunter Löffler, University of Frankfurt, Chair of Banking and Finance, P.O. Box 111932, 60054 Frankfurt (Main), Germany. Tel.: ++49-69-79823143; fax: ++49-69-79822143. *E-mail address:* gloeffler@wiwi.uni-frankfurt.de

Acknowledgements

I wish to thank two anonymous referees whose comments lead to an improvement of the paper. I am also grateful for comments received from Hergen Frerichs, Markus Nöth and participants of the 1999 Autumn Seminar of INQUIRE Europe, the 1999 Symposium on Finance, Banking and Insurance at the University of Karlsruhe and the 2000 Banking Workshop at the University of Münster.

1. Introduction

In the past few years, several new models for the measurement of portfolio credit risk have been proposed (cf. Crouhy, Galai and Mark, 2000). They have the potential to effect major changes in the ways banks are managed and regulated. So far, however, little is known about the reliability of these models. Nickell, Perraudin and Varotto (2001) is the only paper which tests the predictive ability of portfolio models on an out-of-sample basis.

The major reason for the scarcity of research is the lack of data suitable for backtesting (cf. Jackson and Perraudin, 2000). In this paper, I therefore use Monte Carlo simulations to quantify the accuracy of credit risk models. More precisely, I analyse the impact of uncertainty about input parameters on the precision of measures of portfolio risk. I confine the analysis to losses from default, i.e., exclude the risk of credit quality changes, and model default correlations by means of correlated latent variables. The framework builds on CreditMetrics (JPMorgan, 1997), and closely resembles the one used by the Basel Committee on Banking Supervision (2001) to adjust capital requirements for concentration risks.

The necessary inputs for assessing default risk are default rates, recovery rates, and default correlations. They are usually derived from historical data, which means that their precision can be inferred using standard statistical methodology. This is the first step of the analysis in this paper. In the second, I determine the accuracy of Value at Risk (VaR) measures in the presence of noisy input parameters. This is done separately for portfolios which differ in their average credit quality and in diversification across obligors.

The aim of such an analysis is threefold. First, the results are useful for defining the role credit risk models should play in credit portfolio management and bank regulation. Second, modelling parameter uncertainty allows to compute risk measures which take estimation error into account. Since the loss distribution is a nonlinear function of the input parameters, its estimate can be biased even if the parameter estimates are not. To correct such biases, I employ a Bayesian approach and analyse the predictive distribution, which averages the loss distributions pertaining to different but possibly true parameter values.[1] Finally, the analysis helps to identify inputs with a large marginal benefit of increasing input quality.

---

[1] See, for example, Zellner and Chetty (1965) for an application to regression analysis.

The analysis shows that estimation error in input parameters leads to considerable noise in estimated portfolio risk. The confidence bounds for risk measures are so wide that losses which are judged to occur with a probability of 0.3% may actually occur with a probability of 1%. Several observations, however, suggest that available credit risk models can be useful for risk management purposes even though their application is plagued with data problems. The magnitude of estimation error is comparable to a setting in which VaR estimates can be based on a long time series of portfolio losses, and it differs little between perfectly diversified portfolios and small portfolios with 50 obligors. In addition, the bias in conventional VaR figures which results from estimation error is modest. The relative importance of the three input factors for the quality of VaR estimates depends on the portfolio structure and the extremeness of the events under analysis. The impact of correlation uncertainty, for instance, is larger for more extreme events and for riskier portfolios.

Related papers are Jorion (1996) and Butler and Schachter (1998) who argue that market risk measures should be reported with confidence intervals and show how these can be estimated. The methods are not directly applicable to credit risk measurement because they are based on an analysis of changes in portfolio value. For credit portfolios, historical portfolio returns are typically not sufficient for assessing risk. Butler and Schachter (1998) suggest that precision estimates should be taken into account when credit risk models are evaluated for regulatory purposes. This point is supported by the econometric literature on the evaluation of predictions. As shown by West and McCracken (1998), standard statistical inference about predictive ability is misleading if the prediction is based on estimated parameters. As an illustration, I examine the effects of estimation error on a binomial backtest which is based on the number of cases the VaR was exceeded. If estimation error is neglected in the interpretation of the test, one will too often conclude that the structure of a credit risk model is flawed.

Kealhofer, Kwok and Weng (1998) and Nickel, Perraudin and Varotto (2000) examine the problem of estimating probabilities of rating transitions. Gordy (2000a) gives examples on how changes in input parameters affect the VaR. The current study provides several extensions to these papers. It analyses the three inputs to default risk models in a comprehensive way, it distinguishes between systematic and unsystematic estimation risk, and it shows how to adjust the VaR for estimation risk. Lopez and Saidenberg (2000) suggest cross-sectional resampling techniques to make efficient use of available data and to produce measures of forecast accuracy. The Monte Carlo simulations employed here do not use actual data except for quantifying the quality of parameter estimates.

The paper is organised as follows. Section 2 describes the methods used for computing default risk and the assumptions about the magnitude of estimation error. Section 3 presents the simulation results on the accuracy of VaR figures. Section 4 puts the results into perspective and shows how estimation error can bias risk measures and backtests. Section 5 concludes.

2. Methodology

2.1 Modelling portfolio losses

I examine how estimates of the annual loss of loan portfolios are affected by noisy input parameters. Each portfolio consists of N loans of size $1/N$. A loan is worth $1/N$ times the recovery rate in case of default and $1/N$ otherwise. The portfolios are thus homogeneous in terms of the loans' face value; apart from the generalisations of section 3.3, the estimated risk characteristics (default rates, recovery rates, asset correlations) will also be equal across obligors.

The analysis is conducted for four portfolios. They contain either 50 or an infinite number of loans to different obligors, whose credit quality is similar to borrowers rated BBB or B by Standard & Poor's (S&P). The portfolios with an infinite number of obligors approximate the case of large bank portfolios, in which unsystematic risk is eliminated. They are called asymptotic portfolios in the following, while the portfolios with 50 obligors will be called small. The latter portfolios are indeed small relative to typical bank loan portfolios. The number is rather chosen to match portfolio structures of collateralised bond obligations or of mutual funds investing in corporate bonds.

To assess portfolio credit risk, I employ a methodology similar to CreditMetrics[2]. By defining a default rate and a recovery rate for each obligor, the probability and severity of obligor-specific losses is specified. In a next step, default correlations are modelled based on the asset value model of Merton (1974). A firm is assumed to default if its value falls below a critical level defined by the value of liabilities. Correlations of asset values thus translate into default correlations. More precisely, a firm's logarithmic asset value X_i is modelled through a one-factor model:

[2] Cf. JP Morgan (1997) for a general description of CreditMetrics, and Finger (1999) for a description of the analysis of perfectly diversified portfolios.

$$X_i \quad = \quad wZ \ + \ \sqrt{1-w^2}\,\varepsilon_i \tag{1}$$

with Z denoting a common factor, and ε_i the idiosyncratic risk of obligor i. Depending on the realisation of the asset value X_i and the default probability p, a loan is mapped into one of the two possible states. The model is thus similar to a probit model in which events are driven by latent variables.

In CreditMetrics, both Z and ε_i are assumed to follow a standard normal distribution. In this case, the asset correlation is equal to w^2, and default occurs if X_i is below $\Phi^{-1}(p)$ with Φ denoting the cumulative normal distribution function. Gordy (2000a) notes that the assumption of a normally distributed factor is critical, which is why I employ a more general specification. The literature on stock market returns typically documents departures from normality, and proposes discrete mixtures of normal distributions as one alternative specification (e.g. Kon, 1984). I therefore model the distribution of the common factor as a mixture of normal distributions. Specifically, I assume that the common factor Z is drawn from two distributions which have both mean zero but can differ in their variances:

$$Z \quad = \lambda Z_1 + (1-\lambda) Z_2, \quad Z_1 \sim N(0, \sigma^2(Z_1)) \quad Z_2 \sim N(0, \sigma^2(Z_2)), \tag{2}$$

where λ takes on the value 1 with probability γ, and 0 with probability $(1 - \gamma)$. The variance of Z is $\gamma\sigma^2(Z_1) + (1-\gamma)\sigma^2(Z_2)$; a normal distribution obtains by setting $\sigma^2(Z_1)$ equal to $\sigma^2(Z_2)$. In modelling the idiosyncratic component ε_i I follow CreditMetrics and assume it to be a normal variate with mean zero and variance $\sigma^2(\varepsilon_i)$.[3] For a given default probability p and a realisation of Z, the conditional probability that an obligor defaults is then given by

$$q(Z, p, w, \sigma^2(\varepsilon_i)) \quad = \quad \text{Prob}\left(\varepsilon_i \le \frac{d - wZ}{\sqrt{1-w^2}}\right) = \Phi\left(\frac{d - wZ}{\sqrt{1-w^2}\;\sigma(\varepsilon_i)}\right) \tag{3}$$

where d, the default threshold, is the solution to:

$$\gamma\,\text{Prob}\left(wZ_1 + \sqrt{1-w^2}\,\varepsilon_i \le d\right) + (1-\gamma)\,\text{Prob}\left(wZ_2 + \sqrt{1-w^2}\,\varepsilon_i \le d\right) = p \tag{4}$$

With the assumptions on Z and ε_i, equation (4) can be written as

[3] The shape of the distribution of the idiosyncratic component can also affect the distribution of aggregate defaults in a portfolio. The generalisations of section 3.3, however, indicate that the introduction of estimation error about the distribution of ε_i has a small impact on overall uncertainty.

$$\gamma\,\Phi\left(\frac{d}{\sqrt{w^2\sigma^2(Z_1)+(1-w^2)\sigma^2(\varepsilon_i)}}\right)+(1-\gamma)\,\Phi\left(\frac{d}{\sqrt{w^2\sigma^2(Z_1)+(1-w^2)\;\sigma^2(\varepsilon_i)}}\right)=p \qquad (5)$$

In a perfectly diversified homogeneous portfolio, the proportion of obligors that default will always be equal to the conditional probability $q(Z,p,w,\sigma^2(\varepsilon))$. The proportion of defaulted loans is therefore monotonically related to *Z.* This makes it simple to characterise the unconditional distribution of portfolio losses. The conditional loss is $q(Z,p,w,\sigma^2(\varepsilon))\cdot(1-r)$, with *r* being the recovery rate. Let Q_α^Z denote the α quantile of Z. The Value at Risk (VaR) at the percentile level α is given by:

$$\alpha\ \text{VaR}=q\left(Q_\alpha^Z,p,w,\sigma^2(\varepsilon)\right)(1-r)=\Phi\left(\frac{d-wQ_\alpha^Z}{\sqrt{1-w^2}\,\sigma(\varepsilon)}\right)(1-r). \qquad (6)$$

With a probability of α, portfolio losses are larger than the α VaR. For a homogeneous asymptotic portfolio, exact VaR figures can thus be obtained analytically. Any recovery rate risk, i.e. stochastic realisations of the recovery rate given that the mean recovery is *r*, is eliminated in the asymptotic case.

The loss distribution of small portfolios can be determined through Monte Carlo simulations. First, random asset scenarios X_i are drawn according to the factor model (1). Depending on the realisation of the asset value X_i and the default probability p_i, loans are mapped into one of the two possible states. (Default occurs if X_i is below d_i.) Applying mean recovery rates to all defaulted loans and summing over all loans yields the overall portfolio value.[4] The generation of random scenarios is repeated sufficiently often (20,000 times in this paper) to obtain a distribution of portfolio value. The α VaR is calculated as one minus the α quantile of the simulated distribution of portfolio value.

2.2 Modelling parameter uncertainty

To model parameter uncertainty, I assume that risk managers estimate input parameters directly with historical data. Default probabilities are obtained from tabulated default rates, recovery rates from bond prices shortly after

[4] I neglect recovery risk because it is only of second order importance: in the base case (see section 2.2) the VaR of the small B portfolio is 16.13%. Assuming individual recovery rates to be uniformly distributed between zero and two times the base case recovery rate increases the VaR to 16.50%.

default, while the joint distribution of asset values is estimated from observed changes in asset values.[5] Conditional on an assumption about data availability, the quantification of estimation error is straightforward in some cases. In those in which it is not I aim at being conservative, i.e., assume a relatively high level of estimation risk. As it seems impossible to find a specification generally accepted as representative, section 3.3 will present several generalisations.

Estimation risk has a systematic as well as an unsystematic component. The first arises because the estimate of the population average may differ from the true value, the latter is due to the fact that the obligors in a portfolio may differ from the population average. As an example, consider the problem of estimating the default rate of a borrower rated B. If one uses the historical default frequency of B-rated issuers as an estimate, there are two types of error. Due to sampling error or structural changes, the historical average default rate of B-rated issuers may not be equal to the expected default rate of this rating category. Such errors are systematic, they contaminate default rate estimates for all B-rated borrowers. Even if the historical default frequency and the expected average default rate are identical, an individual borrower may have an expected default probability which is higher or lower than the average. This type of error is unsystematic. Its origin may be misclassification of borrowers or loss of information due to the use of broad rating categories which obscure differences between obligors.

Default rates

As a source for default rates, I use the 1999 ratings performance report of Standard & Poor's (1999), which covers the years 1981 to 1998. In this 18-year period, the average default rate for BBB-rated issuers was 0.22%, and 4.82% for B-rated issuers. These figures are only estimates of the true underlying default probabilities. When estimating the standard error of the mean default rate, i.e. the systematic estimation risk, one should allow for the existence of cyclical patterns in realised default rates. As is known from the econometrics literature, neglecting (positive) autocorrelations leads to an underestimation of standard errors (cf. Greene (1993)).

---

5 If asset values are not observable, equity values can be used instead (cf. JP Morgan (1997)). For non-traded firms, asset values can be inferred with a relative valuation approach.

To capture serial correlations, I first estimate the following autoregressive processes for the annual default rates RATE of issuers rated BBB and B, respectively (t-statistics in parentheses):

$$\text{BBB:} \quad RATE_t = \underset{(2.11)}{0.23} + \underset{(0.44)}{0.12}\, RATE_{t-1} - \underset{(-0.34)}{0.09}\, RATE_{t-2}, \quad R^2=0.02 \tag{7}$$

$$\text{B:} \quad RATE_t = \underset{(3.10)}{4.57} + \underset{(2.32)}{0.57}\, RATE_{t-1} - \underset{(-1.93)}{0.46}\, RATE_{t-2}, \quad R^2=0.34 \tag{8}$$

The lag length of two is chosen because it maximises Akaike's information criterion for both rating categories. The regressions point to the existence of serial correlation in the default risk of B-rated issuers. The p-value of the Ljung-Box Q statistic with two lags is 0.106. For the BBB-rated issuers the evidence is weaker (the Ljung-Box p-value is 0.870), but the estimated coefficients have the same sign as their counterparts from the regression of B default rates.

Because of the small sample size, I use a non-parametric bootstrap procedure to derive a distribution for the true mean default rates. I begin by drawing two consecutive annual default rates from the 18 year history. These two rates are the first two observations in the bootstrap time series. The remaining 16 default rates are recursively generated by applying the estimated regression coefficients from above to the previous two default rates and drawing an error term (with replacement) from the residuals of the regression.[6] Averaging the 18 default rates of the bootstrap sample gives a mean default rate which might be true given the history that was observed. The procedure is repeated 20,000 times to yield a default rate distribution.[7]

In the quantification of unsystematic errors, I take recourse to an analysis by Kealhofer, Kwok and Weng (1998). Based on an application of the Merton (1974) model, the authors estimated default probabilities for borrowers rated by S&P. This was done at two different dates. The results suggest that there is significant dispersion in default probabilities of borrowers with the same letter grade. To give an example, the interquartile range of estimated default rates of BBB-rated issuers was 0.48% in June 1991. I model intra-grade dispersion of default risk to match these findings. Specifically, I assume that individual default rates of BBB-rated borrowers follow a beta distribution in

[6] The bootstrapped default probabilities are constrained to be non-negative.

[7] Since the serial correlation is weak, the standard errors do not differ much (around 5%) from those of a conventional bootstrap which draws with replacement from the realised default rates.

the interval [0, 1.5%], with mean 0.22% and standard deviation 0.05%. For B-rated borrowers, I choose a mean of 4.82%, a standard deviation of 3.5% and the interval [0, 25%]. The beta distribution is chosen because of its flexible shape. Table 1 shows the resulting quantiles of the distributions. The beta distribution does not allow a perfect match of the quantiles documented in Kealhofer, Kwok and Weng (1998), but the differences are small. Since estimation risk can be reduced by using finer letter grades or quantitative models for default rate prediction,[8] the assumptions on estimation risk made here can be regarded as conservative. To add systematic estimation risk, the beta distributions are shifted to the right or left such that the mean is equal to a particular bootstrapped true average default rate.[9]

Recovery rates

Estimates of recovery rates can be based on studies of recovery on defaulted bonds. The loans analysed in this paper are assumed to share the properties of senior unsecured bonds. According to a study by Standard & Poor's (1999), which is based on 180 observations, this class has a mean recovery rate of 49.60% with a standard deviation of 26.51%. If an estimate of the mean recovery rate is based on this sample, it will have a standard error of 26.51% / $\sqrt{180}$. As the limiting distribution is normal and the sample size is large (180), I assume the true mean recovery rate to be normally distributed around 49.60%.

In a study of recovery rates, Altman and Kishore (1996) report that, across industry groupings with more than 10 defaults, mean recoveries of senior unsecured bonds vary between 30.83% and 71.91%. This evidence forms the basis for the assumption about unsystematic estimation error. Individual mean recovery rates are assumed to be uniformly distributed within a 40% interval around the true overall mean recovery rate. To add systematic estimation error, the overall mean rate is modelled as described above.

Correlations

Asset correlations are assumed to be driven by a single factor as in (1). To quantify estimation risk, I assume that the parameters of the factor model have been estimated with a sample of 60 returns generated by that model. With monthly observations, this corresponds to a five year history of asset returns, a time span often used in practice to estimate volatilities or correlations. According to the Basel Committee on Banking Supervision

[8] Cf. Moody's (2000) for an example of a default prediction model.

(2001), an average asset correlation of 0.2 is consistent with industry practice. I therefore consider a case where the correlation for each pair of obligors has been estimated to be 0.2, being equivalent to an estimate of 0.2 for w^2, and 1 for the variances of Z and ε.[10] Systematic estimation error results from uncertainty about the distribution of the common factor *Z.* Unsystematic noise stems from erroneous estimates of the factor sensitivities w_i and the idiosyncratic risk $\sigma^2(\varepsilon_i)$.

The common factor *Z* was taken to be a mixture of two normal distributions, which nests the standard assumption of a normally distributed factor. The existing literature on CreditMetrics provides no guidance on how to ascertain the shape of the distribution. I account for this uncertainty by assuming the information on the distribution characteristics to be diffuse. I use a common measure of nonnormality, the kurtosis, to specify the available information. For a given variance of Z, the probability of extreme factor realisations increases with the fatness of the tails, for which the kurtosis provides a measure. I assume that the excess asset kurtosis is set to 1.5, whereas its true value is uniformly distributed in the interval 0 to 3. I restrict it to be positive because the mixture of two normals presented in (4) cannot produce distributions with negative kurtosis. Note that the kurtosis of the S&P 500 index is 0.74 when computed with annual log returns from 1971 to 2000. An interval ranging from 0 to 3 may appear small when compared to the observed kurtosis of daily returns (cf. Kon, 1984). The credit risk analysis, however, is conducted on a longer horizon. Empirically, leptokurtosis tends to decrease for longer return horizons, a finding consistent with the Central Limit Theorem driving returns towards normality. For the S&P 500 data used above, the kurtosis is 31.34 and 3.37 for daily and monthly returns, respectively (see Campbell, Lo and McKinlay, 1997, for further evidence).

The parameters describing *Z* are determined as follows. I set the mixing probability γ equal to 0.5. As will be shown in the generalisations, this turns out to generate the largest estimation error in portfolio losses. For a fixed γ, the variance and the kurtosis specified for *Z* uniquely determine the variances of the two mixing distributions.

---

[9] The simulated default probabilities are constrained to be non-negative.
[10] Gordy (2000a) calibrates the factor sensitivity *w* to match historical default rate volatilities. His estimates for w^2 vary between 0.015 and 0.125. In the generalisations, I also examine the case in which w^2=0.05.

Uncertainty about the variance of Z is modelled by assuming that the estimate is based on 60 observations. I approximate the distribution of the estimate with the chi-squared distribution (60 degrees of freedom), which is exact when Z is normally distributed. Errors in the estimation of the variance and the kurtosis are taken to be independent.[11]

The systematic estimation error works in two ways. For a given common factor sensitivity *w* and a common idiosyncratic variance $\sigma^2(\varepsilon)$, the asset correlation increases with the variance of *Z* according to the following formula:

$$Corr(X_i, X_j) \quad = \quad \frac{Cov(X_i, X_j)}{\sigma(X_i)\sigma(X_j)} = \frac{w^2\, \sigma^2(Z)}{w^2\, \sigma^2(Z) + (1-w^2)\sigma^2(\varepsilon)} \tag{9}$$

Thus, the average asset correlation is subject to estimation error because the variance of *Z* is unknown. Second, for a given variance of *Z*, the probability of extreme factor realisations varies with the kurtosis of Z. Note that noisy estimates of the distribution of *Z* do not imply that the unconditional default rate has been misestimated. In the implementation of the model, the default threshold *d* is determined for a given true default rate *p* and a given true distribution of *Z*.

In modelling unsystematic error, I use the results from OLS-inference stating that $\hat{\sigma}^2(\varepsilon_i)$, the estimate of the variance of ε_i, follows a chi-squared distribution, while an estimate of the factor sensitivity w_i follows a t-distribution.[12] To obtain the distribution of the average correlation under these assumptions, I use Monte Carlo simulations which involve 20,000 random draws of the regression parameters for each obligor.

This concludes the quantification of estimation error in input parameters. Systematic estimation error is due to uncertainty about the average default probability *p*, the average recovery rate *r* and the distribution of the common factor *Z*. Unsystematic error is added through the need to estimate individual default rates p_i, individual recovery rates r_i, individual factor sensitivities w_i and individual idiosyncratic variances $\sigma^2(\varepsilon_i)$. I will refer to the set of parameter estimates described above as the base case parameterisation. The magnitude of the assumed estimation risk is illustrated in Table 2. Along with

---

[11] A simulation analysis shows that, in a sample containing 60 values of Z generated by the case base model, the correlation of the sample variance and the sample kurtosis is 10%.

[12] Cf., for example, Greene (1993).

the base case estimates it contains corresponding 95% confidence intervals for portfolios with 50 or an infinite number of obligors. Uncertainty about default correlations is captured through two figures: the average asset correlation and the kurtosis of the common factor. The confidence intervals are obtained analytically for the systematic error in recovery rates and through simulation (20,000 trials) otherwise. In the case of correlations, the average correlation of a trial is taken to be the correlation of two obligors which have the average parameter values of that trial. An inspection of the confidence intervals for small and asymptotic portfolios shows that the additional noise brought about by unsystematic estimation error is small. It does not immediately follow that unsystematic estimation risk is smaller than systematic one, but unreported analyses confirm this conjecture for each input.

2.3 Computational issues

Once estimation error is introduced, a portfolio which was judged to be homogenous need no longer be, for instance because the factor sensitivities which had been estimated to be uniform across obligors do in fact differ. Due to the nonlinearities involved in the determination of credit risk, heterogeneity can affect portfolio risk. As an example, consider the conditional default probability $q(Z,p,w,\sigma^2(\varepsilon))$ at $Z=-2$, $w^2=0.2$ and $\sigma^2(\varepsilon)=1$. An inspection of the function shows that the conditional default probability is concave in p. In consequence, a homogeneous portfolio with default rate p^* will have a higher conditional default rate than an heterogeneous one where the mean default rate is p^*.

It follows that the risk of asymptotic portfolios is changed by the introduction of unsystematic estimation risk even though the average parameter values are identical to the base case. However, there is no uncertainty about the exact consequences because the distribution of parameters within the portfolio exactly mirrors the distribution of the true parameters assumed in the modelling of unsystematic estimation risk.[13] I neglect this complication in the paper and use equation (6), which yields the VaR of homogeneous portfolios, to determine the risk of the asymptotic portfolios. That is, I ignore the predictable influence of unsystematic estimation error and use parameters which are uniform across borrowers, and thus only subject to systematic estimation risk. The motivation is that I plan to compare the risk

[13] Strictly speaking, this is only true for the case in which there is no uncertainty about the true magnitude of estimation risk.

figures obtained in the base case to ones adjusted for estimation error. The comparison would be invalid if the two risk measures were computed for a homogeneous and a heterogeneous portfolio, respectively.

For small portfolios, risk is determined through Monte Carlo simulations. Both unsystematic and systematic estimation error is modelled, so that portfolios will be heterogeneous once estimation error is introduced. The discussion of the results will account for this differential treatment of homogeneity. In the Monte Carlo simulations, I use a single set of standard normal random numbers for the generation of individual asset scenarios. This reduces simulation error across different specifications of input parameters. Since the loss distributions of small portfolios are discrete, the probability of exceeding the α VaR of a small portfolio need not be exactly equal to α. I determine the VaR such that the probability is α or less.

To assess the effects of estimation risk, I first specify which sources of estimation risk are taken into account. Four variants are considered: default rate uncertainty only, recovery rate uncertainty only, correlation uncertainty only, and the three sources of uncertainty together. I then draw true values of the parameters according to their specified distributions, and use these parameters to calculate measures of portfolio risk. If, for instance, all three sources of uncertainty are modelled, I draw for each obligor a true default rate, a true recovery rate, and true parameters of the one-factor model. Based on these parameters, I calculate risk measures. This procedure is repeated 20,000 times, resulting in a frequency distribution for the risk measures.

In the presence of estimation risk, one cannot be sure that the loss distribution pertaining to the estimated parameter set is the true one, even if the estimates satisfy criteria like unbiasedness or efficiency. The loss distribution should rather be regarded as a mixture of distributions, each having different and possibly true parameter values. In order to take estimation error explicitly into account, one can compute the VaR of the predictive distribution which averages over the unknown parameter values, instead of examining the distribution pertaining to the single parameter vector one has estimated. As an example, consider the regression $y_t=a+bx_t+u_t$, $t=1,..T$, where u_t is normally distributed. Suppose that y_{T+1} is to be predicted based on x_{T+1}, a known value of the regressor. One source of forecast error is sampling error in the parameter estimates. The coefficients a and b as well as the variance of u_t have to be estimated based on observations 1 to T. Additional uncertainty stems from the unknown realisation of the error term

u_{T+1}. The conventional approach in risk management ignores sampling error in the parameters which corresponds to assessing the prediction error based on the estimated variance of u_{T+1}. In the example, the forecast error would thus be assumed to be normally distributed. Taking parameter uncertainty into account results in a multivariate t distribution for the forecast error (Zellner and Chetty, 1965). This predictive density is obtained by integrating with respect to the unknown parameters.

In section 4, I will analyse the predictive distribution of the asymptotic portfolios, taking all three sources of estimation error into account. Recall that the unknown parameters are the unconditional default probability p, the mean recovery rate r, and the variance and kurtosis of Z. The joint distribution of these parameters is simulated using K=20,000 trials. The trials are equally likely to represent the true loss distribution, so that the predictive loss distribution is a simple average of the K loss distributions which pertain to the K simulated parameter combinations.

For an individual parameter combination k, the probability that portfolio losses are greater than L is given by the probability that the conditional default rate $q(Z,p_k,w,\sigma^2(\varepsilon))$, which was introduced in equation (3), is greater than L divided by one minus the recovery rate r_k. Using the predictive distribution, the probability that portfolio losses are greater than L is thus:

$$\begin{aligned}
\text{Prob}(Loss \geq L) &= \frac{1}{K}\sum_{k=1}^{K}\text{Prob}\left(q\,(Z, p_k, w, \sigma^2(\varepsilon)) \geq L/(1-r_k)\right) \\
&= \frac{1}{K}\sum_{k=1}^{K}\text{Prob}\left(\Phi\left(\frac{d_k - wZ}{\sqrt{1-w^2}\,\sigma(\varepsilon)}\right) \geq L/(1-r_k)\right) \\
&= \frac{1}{K}\sum_{k=1}^{K}\text{Prob}\left(Z \leq \frac{d_k - \sqrt{1-w^2}\,\sigma(\varepsilon)\,\Phi^{-1}\left(L/(1-r_k)\right)}{w}\right) \\
&= \frac{1}{K}\sum_{k=1}^{K}\left[\gamma\Phi\left(\frac{d_k - \sqrt{1-w^2}\,\sigma(\varepsilon)\,\Phi^{-1}\left(L/(1-r_k)\right)}{w\sigma_k(Z_1)}\right) + (1-\gamma)\,\Phi\left(\frac{d_k - \sqrt{1-w^2}\,\sigma(\varepsilon)\,\Phi^{-1}\left(L/(1-r_k)\right)}{w\sigma_k(Z_2)}\right)\right]
\end{aligned} \tag{10}$$

where the subscript k shows which parameters differ across the K trials. To obtain a predictive VaR at the percentile level α one only has to find a portfolio loss L for which expression (10) amounts to α. Conversely, if L is chosen to be the conventional VaR at nominal level α, equation (10) yields the corresponding "true" percentile level. The quotation marks shall indicate that some uncertainty remains because the exact magnitude of estimation error is typically not known.

3 Results

3.1 Base case

Table 3 presents the risk characteristics of the different portfolios in the base case, where the asset correlation is 0.2, the mean recovery rate is 49.6%, and default rates are 0.22% (BBB) and 4.82% (B). As measures of risk, three VaR figures are calculated: the 1%, the 5% and the 10% Value at Risk. The numbers for the small portfolio with 50 obligors are subject to simulation error. For the purpose of this paper, it can be neglected because the subject of interest is the additional variation brought about by noisy input parameters.

3.2 Precision of estimated VaR measures

Beginning with the asymptotic portfolios, Figure 1 contains 50% and 95% confidence intervals for the estimated VaR measures, thus illustrating how sure one can be that the base case figures are actually the true ones. The confidence intervals are also reported in Table 4, along with the simulated standard errors of the VaR figures.

In most cases, default rates are the most important source of uncertainty, as measured by the width of the confidence intervals or the standard error. The intuition is that the number of observations available for parameter estimation was assumed to be 18 for default rates compared to 60 and 180 for correlations and recovery rates, respectively. The relative role of correlation uncertainty is larger for more extreme percentile levels, and increases when moving from the BBB to the B portfolio. For the precision of the 1% VaR of the B portfolio, for instance, correlations are the most important source of estimation error. These two patterns can be explained as follows. With rising default rates, the elasticity of default correlations with respect to changes in asset correlations increases (cf. JP Morgan, 1997). For riskier portfolios and for more extreme quantiles, a given error in the asset correlation thus leads to a larger error in the default correlation. There is one exception. In the BBB portfolio, the standard error due to correlation increases when moving from the 5% to the 10% percentile level. It can be explained by the differential effect a change in the kurtosis has on the quantile of the common factor Z. If the true kurtosis is larger than the base case value of 1.5, both the 5% quantile of Z and the default threshold d are shifted away from zero, which means that the increase in risk stemming from a fatter tailed distribution of Z is partly offset. The 10% quantile of Z, by contrast, is (initially) shifted towards zero. In the B case, the differences matter less as the default threshold, being closer to zero, is affected less by the kurtosis of *Z*. Recovery rates, finally, mostly rank third in terms of importance.

The magnitude of the estimation error shall be illustrated with the following examples. The upper bound of the 95% confidence interval of the 1% VaR of the B portfolio is 21.33%. This corresponds to the 0.27% VaR of the distribution assumed in the base case. In other words, an event that is judged to be of probability 0.3% may actually have a probability of 1% or more. For the BBB portfolio, it is the 0.37% VaR of the base case which corresponds to the upper confidence bound of the 1% VaR.

The results for portfolios consisting of 50 obligors, which are shown in Figure 2 and Table 5, are similar to the case of perfect diversification in that default probabilities are the main driver of noise in VaR. The influence of recovery rate uncertainty has increased in relative importance. Due to the discrete nature of the portfolio distribution, the VaR does not necessarily change if default rates or correlations change. Recovery rates, on the other hand, always affect the VaR whenever it is above zero.

The discreteness of the loss distribution is also responsible for a result which may seem counterintuitive. The standard error of the 10% VaR of the BBB portfolio goes down when correlation and recovery rate uncertainty are added to the estimation error in default rates. The modelling of recovery rate uncertainty mitigates discreteness, which can shift VaR measures in both directions. Here, the VaR is reduced on average, which also reduces its standard error.

Another puzzling result is that the standard errors of the VaR figures are sometimes smaller than the corresponding standard errors in the asymptotic case. Again, this can be the result of discreteness problems, especially if recovery rate uncertainty is not modelled. However, there is another effect at work. As described in section 2.3, the heterogeneity produced by unsystematic estimation errors is suppressed in the analysis of asymptotic portfolios, but not in the analysis of small portfolios.

To gauge the effects of this differential treatment of unsystematic estimation risk, I impose homogeneity on the small portfolios. In each run of the Monte Carlo simulations, I set the parameters of each obligor equal to the average value of the parameters drawn for this particular trial. The common default rate taken in trial *k,* for instance, is the arithmetic average of the 50 individual default rates p_{ik}. Unsystematic estimation errors still affect the results as they need not sum to zero across the 50 borrowers of the portfolio. The modification leads to a moderate increase of standard errors. For the 1% VaR, the standard errors now amount to 0.54% and 2.81% for the B and BBB portfolios, respectively, if all three sources of estimation error are accounted for. A similar picture emerges for the 5% and 10% VaR measures.[14]

These results resolve the apparent inconsistencies between Tables 4 and 5, but they do not change the conclusion that there is no significant difference in estimation errors between small and asymptotic portfolios. Regardless of the treatment of homogeneity, confidence bounds and standard errors for the VaR of small portfolios are similar to the corresponding ones for the VaR of asymptotic portfolios. There are two explanations for the result. First, diversification reduces estimation error within the portfolio. Second, the marginal impact of unsystematic risk is small because systematic risk is relatively large. One cannot conclude from the results that unsystematic errors are negligible per se. Ignoring systematic errors in the simulations, the overall standard error of the 1% VaR of the small B portfolio is 1.27%. This is much less than the corresponding standard error in the asymptotic case (2.62%), where only systematic risk matters, but still sizeable.

The result is important because practical applications of credit risk models often involve approximations which obscure individual differences between borrowers. For example, available credit risk models often classify obligors into discrete rating categories. The simulations suggest that such approximations add little noise to measures of portfolio risk if the holdings of a portfolio are not dominated by individual borrowers, industries or other groups of borrowers with correlated estimation errors.

3.3 Generalisations

Heterogeneous portfolios

The portfolios analysed in the previous section were homogeneous in terms of estimated credit qualities and correlations. To check whether this assumption is crucial for the results, I examine portfolios which are evenly split between loans rated BBB and B. Within both groups, I also vary the correlation parameter *w*. Instead of choosing a uniform asset correlation of

[14] Another possibly puzzling result is that VaR measures tend to be quite low in the presence of default rate uncertainty. This is due to the fact that heterogeneity in default rates has a bigger impact on portfolio risk than heterogeneous correlations or recovery rates.

0.2, I set w^2 equal to 0.1 for one half of the borrowers, and 0.3 for the other.[15] The generalisation does not change the important result that unsystematic estimation errors add little to systematic ones. The standard error of the 1% VaR of an heterogeneous portfolio with 50 obligors is 1.51%; in the asymptotic case, the corresponding figure is 1.39%.

Conditioning default probabilities

The default rates used in the calculation of portfolio credit risk were unconditional. In practical applications, default rates are sometimes conditioned on macroeconomic variables. I therefore examine the estimation error arising in this context. I regress the annual default rates of B rated issuers on the high-yield spread from the end of the preceding year, which is a variable often included in default rate prediction models (e.g. Kim, 1999). Here, the spread is calculated as the difference between the redemption yields of the M.Lynch US High Yield Master and the M.Lynch US Government Master indices; data for both indices is available from 1984 onwards.[16] Regressing the 1985-1998 default rates on the yield spread produces an R^2 of 0.59. The estimation error associated with a default rate forecast is captured by the standard error of the OLS prediction; it depends on the value of the conditioning variable. For the 1985-1998 period, the median standard error is 0.63%, which is almost identical to the bootstrapped standard error of the unconditional default rates (0.62%). The example thus suggests that the use of unconditional default rates is not critical. There is another reason why the result is comforting. One may question the reliability of the bootstrap procedure on the grounds that the sampling period 1981-1998 is dominated by a long economic expansion. If this introduces a bias into unconditional default rates, a natural solution is to use conditional default rates. Based on the example, one can expect the associated estimation error to be similar to the one assumed in the base case.

Asset value distribution

Among the assumptions on estimation error the ones pertaining to the asset value distribution may appear to be the most arbitrary. I thus present several modifications:

[15] Within the small portfolios, I set w^2 equal to 0.1 (0.3) for 12 (13) out of the 25 obligors of each rating category.

[16] The index data are obtained from Datastream.

(i) if the kurtosis of the common factor is known to be 1.5, the standard error of the 1% VaR of the asymptotic B portfolio which is due to correlation uncertainty falls from 2.20% to 1.60%.

(ii) if the base case is modified by setting the mixing parameter γ to 0.8 (0.9) instead of 0.5, the standard error of the 1% VaR of the asymptotic B portfolio which is due to correlation uncertainty amounts to 2.01% (1.60%) instead of 2.20%.

(iii) I increase uncertainty by assuming that the true value of the kurtosis is uniformly distributed in the interval [0, 6]. In contrast to the base case, the mixing probability γ is not assumed to be known, but randomly drawn from the interval [0.67, 1). The restriction to this interval is necessary because a γ smaller than 0.67 (and greater than 0.33) cannot produce a kurtosis of 6. The overall standard error of the 1% VaR rises to 0.57% (asymptotic BBB portfolio) and 3.39% (asymptotic B portfolio). The associated 95% confidence intervals increase to [0.71%, 2.77%] and [10.80%, 23.80%], respectively.

(iv) the idiosyncratic risk ε_i was assumed to follow a normal distribution. Similar to the treatment of the common factor *Z*, I introduce estimation error by modelling it as an even mixture of two normal distributions. Again, I assume that the kurtosis can range from 0 to 3. The standard error of the 1% VaR of the small B portfolio which stems from correlation uncertainty remains practically unchanged (2.11% versus 2.12%).

(v) I change the asset correlation parameter w^2 from 0.2 to 0.05. This reduces portfolio risk, but the general picture remains largely the same. In the asymptotic B portfolio, correlation uncertainty is still the most important factor for uncertainty in the 1% VaR, while the other two quantiles are affected primarily by default rate uncertainty. In the asymptotic BBB portfolio, default rate uncertainty matters most in all three cases.

Overall precision of parameter estimates

To illustrate the effects of a change in input quality, I simulate the distribution of the 1% VaR of the asymptotic portfolios assuming that the number of observations available for parameter estimation is twice that used in the base case. There are thus 360 observed defaults to estimate recovery rates, and 120 asset returns for estimating correlations. The range assumed for the true kurtosis is not changed. The precision of the mean default rate is modelled by drawing bootstrap samples of size 36 from the 18-year history of default rates. Table 6 compares the 95% confidence intervals arising in this context with the previous ones reported in Table 4. The relative increase in precision matches the relative contribution to estimation error observed in Table 4. While improvement of default rates matters most when estimating the risk of the BBB portfolio, it ranks only second in the case of the B portfolio.

Such an analysis can be useful when risk managers are to prioritise actions aimed at improving the quality of input parameters. It needs to be complemented by an assessment of the costs of this improvement. Recommendations drawn from a cost-benefit analysis will depend on data availability, portfolio structures, risk management objectives and other factors. Thus, I only present an example: if *(i)* the portfolio can be approximated by the asymptotic BBB portfolio, *(ii)* the base case represents the status quo parameterisation, *(iii)* doubling the number of observations for parameter estimation entails similar costs across inputs, and *(iv)* risk management is based on the VaR due to default, then the analysis suggests to focus on improving estimates of default rates rather than those of correlations or recovery rates.

Correlated estimation errors

The three sources of uncertainty have been assumed to be independent. In practice, recovery rates can be subject to systematic risk (cf. Gordy, 2000b), implying that estimations errors in mean default rates and mean recovery rates could be correlated. To check how such correlations would affect the result, I assume estimation errors in mean default rates and mean recovery rates to be perfectly negatively correlated. This is implemented by pairing the highest simulated true default rate with the lowest simulated true recovery rate, and so forth. Even with such an extreme assumption, the increase in estimation error is moderate. The overall standard error of the 1% VaR of the asymptotic B portfolio, for instance, increases from 2.62% to 2.90%.

Credit Migrations

Finally, I extend the analysis by allowing for credit migrations. I follow the standard CreditMetrics approach which models migrations to any of the eight rating categories (including default). I analyse an asymptotic portfolio which contains BBB-rated bonds. Conditional on the future rating the one-year ahead values of an individual bond are assumed to be as follows (the values are taken from an example in JP Morgan (1997), p.10):

AAA	AA	A	BBB	BB	B	CCC
109.37	109.19	108.66	107.55	102.02	98.1	83.64

Recovery rates and asset correlations are modelled as in the base case, while transition probabilities are taken from Standard & Poor's (1999). Uncertainty about these probabilities is modelled through a bootstrap. One bootstrap sample is constructed through 18 draws (with replacement) from the annual transition probabilities reported in Standard & Poor's (1999).[17]

The 1% VaR of the portfolio is 3.02%, which is a marked increase against the 1.51% of the base case. If all sources of estimation error are modelled, the standard error of this VaR is 0.83%. The magnitude of estimation error is similar to the base case in which default was the only source of risk. There, the standard error of the 1% VaR divided by its estimate was 0.30; allowing for credit migrations, the corresponding figure is 0.27. There is, however, a shift in the relative importance of the various input parameters. Once migration risk is modelled, correlations are more important than probabilities of credit quality changes. While correlations lead to a standard error of 0.70% in the 1% VaR, the standard error due to transition rate uncertainty amounts to only 0.45%. The intuition is that the sensitivity of credit quality correlations to asset correlations increases with the probability of credit quality changes. In the base case, the only possible credit quality change was default, which occurs with a low probability. In the extended model, negative credit quality changes have a probability of more than 5%.

4. Assessing the magnitude and consequences of estimation error

The large confidence intervals for the VaR presented in the previous section could raise doubts about the usefulness of a credit risk analysis. In the following, I will give some information which puts the magnitude of estimation risk into perspective and clarifies its consequences.

In the measurement of market risk, VaR is often determined through historical simulation, i.e., by examining losses that would have occurred in the past under the current portfolio structure. Regulatory authorities consider 250 loss scenarios to be sufficient for measuring risk through historical simulation. Due to differences in forecast horizons and loss distributions, market and credit risk models are difficult to compare. Nevertheless, it is interesting to study the estimation error that would arise if credit risk could be determined through historical simulation.

[17] This procedure neglects serial correlation in transition probabilities, but the previous analysis has indicated that there is little correlation in default rates of BBB issuers.

As in Kupiec (1995), I generate T independent draws from a known probability distribution and take the 1% VaR to be the 1% quantile of the scenarios.[18] This procedure is repeated 20,000 times to construct a sampling distribution for the 1% Var. If the underlying probability distributions is the base case distribution of the asymptotic B portfolio, the standard error of the 1% VaR is 3.13% and 2.21% when the historical simulation consists of 250 and 500 scenarios, respectively. Recall that the assumptions about parameter uncertainty lead to a standard error of 2.62% (cf. Table 4). For the asymptotic BBB case, the simulated standard error of the 1% VaR is 0.75% and 0.45% when the historical simulation comprises 250 and 500 scenarios, respectively. The corresponding figure from Table 4 is 0.45%. The generalisation which doubled uncertainty about the kurtosis of Z lead to a standard error of 0.57%. The results caution against deeming credit risk models to be too inaccurate solely because the available number of observations is small compared to a market risk analysis.

Consider next a bank who wants to combine a riskless government bond portfolio with a perfectly diversified B-rated loan portfolio such that the 1% VaR due to default is 5%. In addition, the bank would like to have a confidence of 75% that the true 1% VaR is at or below 5%. Therefore, the bank should base its decision on a VaR of 17.47% (i.e. the 75% quantile of the estimated VaR) for the risky part of the portfolio, not on the 15.27% of the base case. This difference has only a moderate impact on the admissible portfolio structure. With the base case VaR of 15.27% for a B portfolio, the maximum allowable weight for the B loans is 32.7%. Applying the more conservative estimate reduces the admissible weight to 28.6%. If the institution would like to be 97.5% confident to have a portfolio with a 1% VaR of 5%, it could invest 23.4% of its portfolio into B loans.

The estimated VaR may understate risk, but it may also overstate it. An approach which takes both sides into account is to treat estimation error as an additional component of risk and assess its overall effects on the distribution of portfolio value. This can be done by modelling the portfolio distribution as a mixture of distributions as described in Section 2.3. For the asymptotic portfolios, Table 7 presents the conventional VaR computed with the base case parameters as well as predictive VaRs computed over the

[18] For T=250, I calculate the 1% VaR as the midpoint between the 2nd and the 3rd largest loss.

20,000 simulated distributions which formed the basis of the results presented in Table 4.

The results show that the conventional VaR underestimates the predictive VaR which takes estimation error into account. The magnitude of the bias ranges from of 3 to 69 basis points. Another way of comparing conventional and predictive VaR measures is to compute the percentile level of the standard VaR under the predictive distribution. For the 1% VaR, this predictive percentile level is 1.11% and 1.15% for the BBB and B portfolio, respectively. Based on the assumptions about estimation risk, the best guess for the percentile level of the base case VaR figures (at the nominal level of 1%) would thus be 1.11% and 1.15%, respectively. I compute the same figures for one of the generalisations, in which the uncertainty about the kurtosis was doubled (see section 3.3). In the computation of the nominal 1% VaR the parameter γ is set to its expected value, which is 0.835 because it is randomly drawn from the interval [0.67, 1). The predictive percentile levels for the nominal 1% VaR are 0.97% (BBB) and 0.98% (B), showing that the bias can be in either direction. It may appear counterintuitive that an increase of uncertainty leads to a decrease of bias. The base case and the generalisation, however, are not directly comparable because the parameters for the calculation of the nominal VaR are not identical, and because γ is assumed to be known in the base case.

The documented biases appear to be modest. For the setting analysed here, conventional VaR figures can be regarded as reasonable approximations to the true risk of a portfolio. However, it is conceivable that estimation error adjustments can be economically important, for instance if risk management focuses on more extreme events than the 1% quantile. As an example, I calculated the predictive percentile level of the conventional 0.1% VaR in the base case setting. In the asymptotic B case, it turns out to be 0.15%, fifty percent larger than the assumed percentile level.

It is also interesting to compare the documented estimation errors to the quality with which credit risk models track the risks of actual portfolios. Nickell, Perraudin and Varotto (2001) calculated 1% VaRs using the CreditMetrics approach for randomly selected portfolios of 10 bonds over the years 1988 to 1998. The VaR was violated several times by a subset of portfolios, and once by most of the portfolios. On that occasion, the performance of 10% of the portfolios was 200 basis points or more below the VaR. This evidence could lead to the conclusion that the CreditMetrics model itself is inappropriate for the portfolios analysed. Another explanation would be that the model, while being appropriate, lead to misleading results because of noisy input parameters. The analysis of this paper is not directly comparable to Nickell, Perraudin and Varotto (2001) because they also model the risk of rating changes other than default, and because portfolios differ from the ones analysed here. However, the uncertainty about the 1% VaR of the portfolio containing 50 B loans suggests that underestimating the VaR by 200 basis points is not an infrequent event.

In a regression framework, West and McCracken (1998) show that standard tests of predictive ability are biased in the presence of estimation error. Even if estimates make efficient use of available information, estimation error negatively affects the forecast performance. The potential effects on backtests of credit risk models shall be illustrated with some examples. Assume that there are M independent realisations of annual credit losses. A test can be based on V, the number of cases in which portfolio losses were larger than the α VaR. The observed number of such exceptions follows a binomial distribution. If V is larger than what is to be expected by chance, one can conclude that the risk model has underestimated the VaR.

For the asymptotic portfolios, a backtest can account for estimation errors as follows. Insert the base case α VaR into equation (10). For each of the K simulated true parameter sets, (10) gives the true percentile level α_k of the estimated α VaR. Let $B(V,\alpha,M)$ denote the binomial probability of observing V exceptions or more if there are M independent draws and the true probability of an exception is α. Since the K parameter sets are equally likely, the overall probability of observing V or more exceptions under the predictive distribution is given by:

$$\text{Prob}(\mathit{exceptions} \geq V) = \frac{1}{K}\sum_{k=1}^{K} B\left(V, \alpha_{\mathrm{k}}, M\right) \tag{11}$$

If estimation error is ignored, the assumed probability of observing V or more exceptions of the α VaR is $B(V,\alpha, M)$. For some selected values of M and V, Table 8 lists the probabilities of observing V or more exceptions of the 1% VaR. With 10 observations, the probability of observing two or more exceptions of the 1% VaR of the B portfolio increases from 0.43% to 0.69% once estimation error is taken into account. This bias is small when compared to the tests based on 100 or 1000 observations. With M=1000, exceptions that were judged to occur with a probability of 0.69% may have a probability of 15.08%. The explanation is that the sampling error in the observed number of exceptions goes down as the number of observations increases. In consequence, the ratio of estimation error to sampling error rises, and so do biases in estimated probabilities.

One may argue that the number of observations available for backtests of credit risk models will typically be small, maybe as small as ten. The results would then suggest that estimation error is of little importance. However, it is conceivable that the power of backtests can be increased, for example by analysing monthly instead of annual losses, or by exploiting cross-sectional information as in the proposal of Lopez and Saidenberg (2000).

Whether estimation error should be taken into account also depends on the purpose of a test. If the aim is to assess whether the prediction itself was correct, it seems appropriate to ignore it. The case is different if backtests are used to judge whether the structure of a model is adequate or whether a risk management system was set up diligently. If risk managers make efficient use of available information, they should not be blamed for prediction errors which stem from unavoidable estimation errors.

5. Conclusion

The analyses have revealed that estimates of portfolio credit risk are very sensitive to uncertainty about input parameters. This statement should not be interpreted as a negative verdict on available credit risk models. The confidence bounds for the credit VaR are similar to those that would obtain if the VaR could be estimated with 250 observations on portfolio losses, a sample size which regulators consider sufficient in a market risk setting. Another result which supports the usefulness of credit risk models is that unsystematic estimation risk adds little to systematic one. Applications of credit risk models sometimes rely on relatively crude classifications of individual borrowers. Even in a small portfolio with 50 obligors, such simplifications do not seem to produce much additional noise.

Even if parameter estimates are unbiased, conventional VaR estimates need not be. Biases can be corrected by analysing the predictive loss distribution which averages over the unknown parameter values. At least in the cases considered in the paper, biases in conventional VaR estimates are small. This is a positive result because available implementations of credit risk models do not account for estimation error. Nevertheless, risk managers and regulators should be aware of the issue. There may be situations in which a correction of biases is important. In addition, estimation errors are relevant for interpreting the results of backtests. Estimation risk can dramatically increase the probability of observing portfolio losses which exceed the VaR.

Evaluators who ignore parameter uncertainty could be lead to conclude that the structure of a model is flawed even though its poor performance is solely due to unavoidable noise in inputs.

Another question risk managers may face is how much to invest in the improvement of parameter estimates and on which of the input parameters to concentrate. General recommendations are difficult to derive because they crucially depend on data availability, portfolio structure and risk management objectives. The contribution of the paper is that it proposes a framework for assessing the benefits of increasing parameter accuracy.

References

Altman, E., and V. Kishore, 1996, Almost everything you wanted to know about recoveries on defaulted bonds, Financial Analysts Journal 52, 57-64.

Basel Committee on Banking Supervision, 2000, Credit ratings and complementary sources of credit quality information, Basel.

Basel Committee on Banking Supervision, 2001, Overview of the new Basel capital accord, Basel.

Butler, J.S. and B. Schachter, 1998, Estimating value at risk with a precision measure by combining kernel estimation with historical simulation. Review of Derivatives Research 1, 371-390.

Campbell, J., A. Lo and C MacKinlay, 1997, The econometrics of financial markets, Princeton University Press, Princeton.

Crouhy, M., D. Galai and R. Mark, 2000, A comparative analysis of current credit risk models, Journal of Banking and Finance 24, 59-117.

Finger, C., 1999, Conditional approaches for CreditMetrics portfolio distributions, CreditMetrics Monitor, April, 14-33.

Gordy, M., 2000a, A comparative anatomy of credit risk models, Journal of Banking and Finance 24, 119-149.

Gordy, M., 2000b, A risk-factor model foundation for ratings-based bank capital rules, working paper, Board of Governors of the Federal Reserve System.

Greene, W., 1993, Econometric Analysis, 2nd Edition, New York.

Jackson, P., and W. Perraudin, 2000, Regulatory implications of credit risk modelling, Journal of Banking and Finance 24, 1-14.

Jorion, P., 1996, Risk^2: measuring the risk in value at risk, Financial Analysts Journal 52, 47-56.

JP Morgan, 1997, CreditMetrics – Technical document.

Kealhofer, S., S. Kwok and W. Weng, 1998, Uses and abuses of default rates, CreditMetrics Monitor, 1st Quarter, 37-55.

Kim, J., 1999, A way to condition the transition matrix on wind. Working paper, Riskmetrics Group.

Kon, S., 1984, Models of stock returns - a comparison, Journal of Finance, 39, 147-165.

Kupiec, P., 1995, Techniques for verifying the accuracy of risk measurement models. Journal of Derivatives, 73-84.

Lopez, J., and M. Saidenberg, 2000, Evaluating credit risk models, Journal of Banking and Finance 24, 151-165.

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

Moody's Investor Service, 2000, Moody's public firm risk model: a hybrid approach to modeling short term default risk.

Nickel, P., Perraudin, W., Varotto, S., 2001, Ratings- versus equity-based credit risk modeling: An empirical analysis. Working paper, Bank of England, Birkbeck College.

Nickel, P., W. Perraudin, and S. Varotto, 2000, Stability of rating transitions, Journal of Banking and Finance 24, 203-227.

Standard & Poor's, 1999, Ratings performance 1998. Stability and Transition.

West, K.D. and M.W. McCracken, 1998, Regression-based tests of predictive ability, International Economic Review 39, 817-840.

Zellner, A. and V. Chetty, 1965, Prediction and decision problems in regression models from the Bayesian point of view, Journal of the American Statistical Association 60, 608-616.

Table 1
Intra-grade dispersion of default probabilities

	Rating	Quantiles of default probabilities divided by their mean			
		10%	25%	75%	90%
June 1991 from KKW (1998)[a]	BBB	0.15	0.22	1.11	2.41
June 1995 from KKW (1998)	BBB	0.17	0.30	1.10	2.37
Assumption in this paper	BBB	0.10	0.29	1.46	2.30
June 1991 from KKW (1998)	B	0.21	0.39	1.08	2.11
June 1995 from KKW (1998)	B	0.13	0.26	1.15	2.15
Assumption in this paper	B	0.20	0.43	1.42	2.03

[a]KKW (1998) is an empirical study by Kealhofer, Kwok and Weng who use the Merton model to estimate default probabilities for borrowers rated by S&P.

Table 2
Summary of input parameters and the assumed magnitude of estimation error (in percent except for kurtosis)

	Parameter	Estimate	95% confidence interval			
			50 obligors		∞ obligors	
Mean default rate (BBB)	p	0.22	0.11	0.37	0.13	0.36
Mean default rate (B)	p	4.82	3.58	6.70	3.92	6.36
Mean recovery rate	r	49.60	44.58	54.62	45.73	53.47
Average asset correlation [a]		20.00	14.39	27.29	15.26	27.04
Kurtosis of common factor Z [b]		1.50	0.075	2.925	0.075	2.925

[a] Affected by factor sensitivities w (w^2 = 0.2 in base case), the factor variance $\sigma^2(Z)$ and idiosyncratic variances $\sigma^2(\varepsilon_i)$. These variances are set to one in the base case.

[b] Determines the variances ($\sigma^2(Z_1)$, $\sigma^2(Z_1)$) of the two normal distributions which describe the factor Z. The mixing probability γ is fixed at 0.5.

Table 3
Portfolio distribution in the base case without estimation error (in percent of portfolio value)

Portfolio	1% VaR[a]	5% VaR	10% VaR
50 loans (BBB)	2.02	1.01	0.00
50 loans (B)	16.13	8.06	6.05
∞ loans (BBB)	1.51	0.43	0.21
∞ loans (B)	15.27	7.79	5.13

[a] With a probability of α, portfolio losses are larger than the α VaR

Table 4

Simulated distribution of the VaR due to default of asymptotic portfolios[a]

Source of estimation error	VaR of a portfolio with an infinite number of BBB obligors					VaR of a portfolio with an infinite number of B obligors				
		Quantiles					Quantiles			
	Std.error	2.5%	25%	75%	97.5%	Std.error	2.5%	25%	75%	97.5%
Panel A: 1% VaR										
Recovery rate	0.059	1.40	1.47	1.55	1.63	0.596	14.08	14.84	15.65	16.42
Correlation	0.266	1.03	1.30	1.69	2.03	2.202	11.56	13.72	16.76	20.04
Default rate	0.341	0.93	1.39	1.85	2.27	1.201	13.38	14.91	16.54	18.10
All	0.453	0.83	1.29	1.89	2.59	2.616	11.19	13.91	17.47	21.33
Panel B: 5% VaR										
Recovery rate	0.017	0.40	0.42	0.45	0.47	0.304	7.18	7.57	7.98	8.38
Correlation	0.015	0.41	0.43	0.45	0.47	0.661	6.88	7.51	8.39	9.46
Default rate	0.118	0.24	0.39	0.55	0.71	0.829	6.54	7.55	8.67	9.80
All	0.124	0.25	0.40	0.56	0.73	1.144	6.28	7.52	9.04	10.77
Panel C: 10% VaR										
Recovery rate	0.008	0.19	0.21	0.22	0.23	0.200	4.73	4.98	5.25	5.51
Correlation	0.030	0.16	0.19	0.24	0.27	0.263	4.82	5.06	5.42	5.83
Default rate	0.063	0.11	0.19	0.27	0.36	0.624	4.21	4.95	5.79	6.66
All	0.073	0.11	0.19	0.28	0.39	0.733	4.18	5.00	5.98	7.06

[a] The results show the distribution of the portfolio Value at Risk (in percent of portfolio value) in the presence of estimation risk. Estimation error in the following input parameters is modelled: Default rates (estimates based on the S&P history), recovery rates (estimates based on prices of defaulted bonds), default correlations (estimates based on joint distribution of asset values).

Table 5

Simulated distribution of the VaR due to default of small portfolios with 50 obligors [a]

Source of Estimation error	VaR of a portfolio with 50 BBB obligors					VaR of a portfolio with 50 B obligors				
		Quantiles					Quantiles			
	Std.error.	2.5%	25%	75%	97.5%	Std.error	2.5%	25%	75%	97.5%
Panel A: 1% VaR										
Recovery rate	0.110	1.89	2.03	2.18	2.33	0.822	14.64	15.71	16.81	17.88
Correlation	0.026	2.02	2.02	2.02	2.02	2.121	13.10	15.12	18.14	21.17
Default rate	0.446	1.01	2.02	2.02	3.02	1.725	12.10	14.11	16.13	18.14
All	0.450	1.32	1.89	2.41	3.11	2.648	10.80	13.63	17.17	21.08
Panel B: 5% VaR										
Recovery rate	0.068	0.85	0.93	1.03	1.11	0.427	7.66	8.21	8.79	9.34
Correlation	0.007	1.01	1.01	1.01	1.01	0.660	8.06	8.06	9.07	10.08
Default rate	0.137	1.01	1.01	1.01	1.01	1.130	6.05	8.06	9.07	10.13
All	0.210	0.62	0.91	1.14	1.29	1.334	6.19	7.68	9.45	11.41
Panel C: 10% VaR										
Recovery rate	0.000	0.00	0.00	0.00	0.00	0.290	5.25	5.63	6.02	6.39
Correlation	0.000	0.00	0.00	0.00	0.00	0.412	5.04	6.05	6.05	6.05
Default rate	0.451	0.00	0.00	1.01	1.01	0.862	4.03	5.04	6.05	8.06
All	0.339	0.00	0.00	0.63	0.89	0.911	4.32	5.37	6.58	7.87

[a] The results show the distribution of the portfolio Value at Risk (in percent of portfolio value) in the presence of estimation risk. Estimation error in the following input parameters is modelled: Default rates (estimates based on the S&P history), recovery rates (estimates based on prices of defaulted bonds), default correlations (estimates based on joint distribution of asset values).

Table 6

Effects of doubling the number of observations for parameter estimation on the confidence interval for the 1% VaR of asymptotic portfolios[a]

Portfolio	Source of estimation error	95% confidence interval for the 1% VaR				Decrease of interval width
		base case		after doubling the number of observations		
BBB	Recovery	1.40	1.63	1.41	1.57	0.07
	Correlation	1.03	2.03	1.05	1.93	0.13
	Default rate	0.93	2.27	1.25	2.15	0.43
B	Recovery	14.08	16.42	14.17	15.84	0.67
	Correlation	11.56	20.04	11.86	18.86	1.47
	Default rate	13.38	18.10	14.20	17.48	1.44

[a] In the base case, default rates are estimated with the S&P history (18 years), recovery rates are estimated using prices of defaulted bonds (180 observations), and asset correlations are estimated based on 60 observations. In the second case, the number of observations available for estimation is doubled for each parameter.

Table 7

Biases of conventional VaR estimates in the presence of estimation risk (in percent of portfolio value)[a]

Portfolio	Percentile level	Standard VaR	Predictive VaR	Difference
BBB	1%	1.51	1.60	0.09
	5%	0.43	0.48	0.05
	10%	0.21	0.24	0.03
B	1%	15.27	15.96	0.69
	5%	7.79	8.29	0.50
	10%	5.13	5.54	0.41

[a] The Value at Risk of homogeneous loan portfolios with an infinite number of obligors is computed in two different ways. The standard VaR is based on one set of estimated input parameters. The predictive VaR is computed over 20,000 distributions of portfolio value which are equally likely under the assumptions about estimation risk in input parameters.

Table 8

Probabilities for exceptions of the 1% VaR of asymptotic portfolios (in percent)[a]

			Prob (exceptions ≥ *V*)	
Portfolio	*Observations M*	*V*	No estimation error	With estimation error
BBB	10	2	0.43	0.61
	100	5	0.34	1.16
	1000	19	0.69	11.21
B	10	2	0.43	0.69
	100	5	0.34	1.66
	1000	19	0.69	15.08

[a] The table lists probabilities of observing *V* or more exceptions of the 1% VaR if there are *M* independent observations on portfolio losses. Without estimation error, the probabilities follow directly from the binomial distribution. To model estimation error, the binomial probabilities are averaged over the possibly true percentile levels of the estimated VaR.

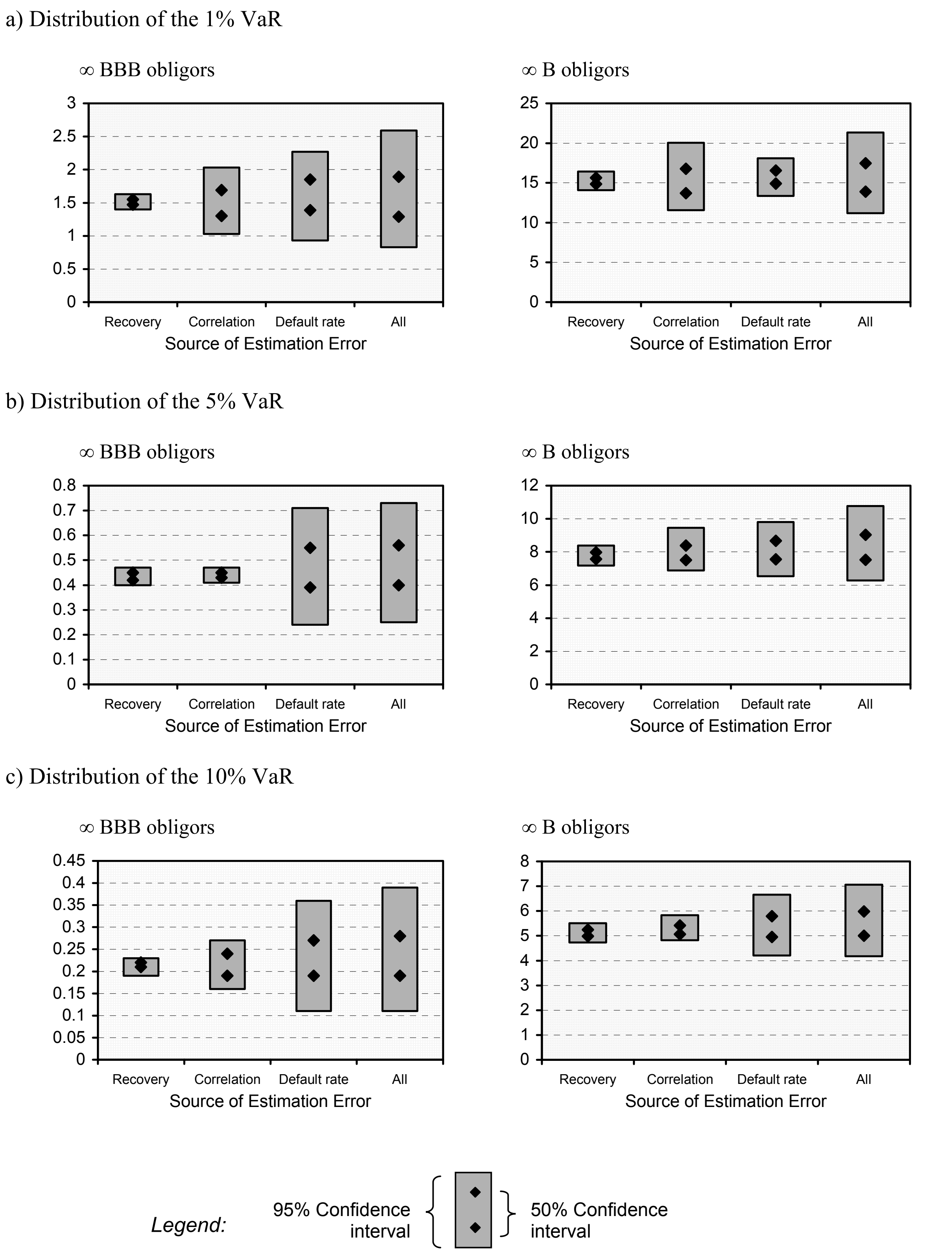

Fig. 1. Simulated distribution of the VaR due to default of asymptotic portfolios

Notes: The figure shows confidence intervals for the portfolio Value at Risk (in percent of portfolio value) in the presence of estimation risk. Estimation error in the following input parameters is explicitly modelled: Default rates (estimates based on the S&P history), recovery rates (estimates based on prices of defaulted bonds), default correlations (estimates based on joint distribution of asset values).

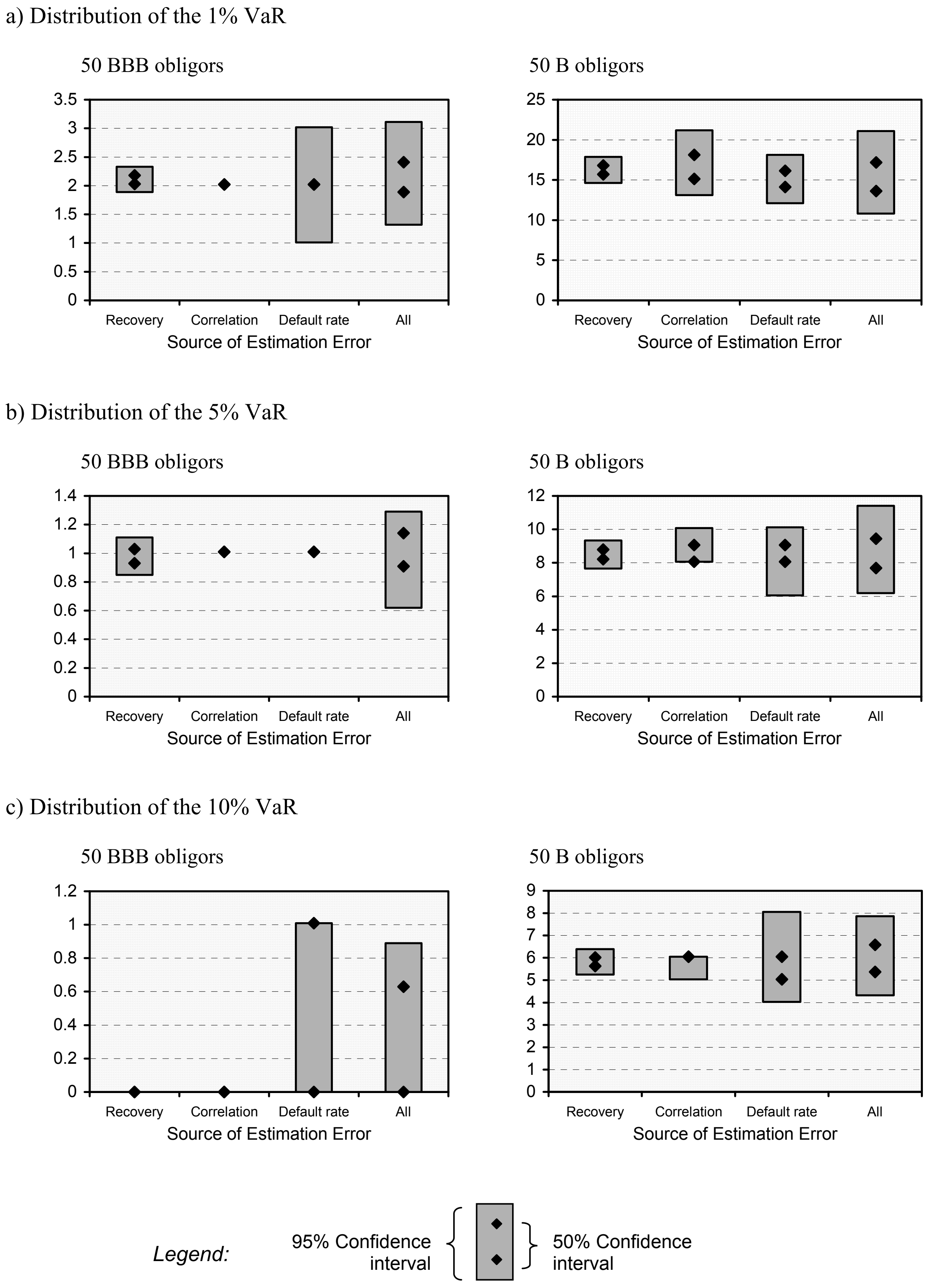

Fig. 2. Simulated distribution of the VaR due to default of small portfolios with 50 obligors
Notes: see Figure 1.