DEFAULT CORRELATIONS AND THE EFFECT OF ESTIMATION ERRORS ON RISK FIGURES

CHRISTINA R. NIETHAMMER AND LUDGER OVERBECK

Abstract. In credit risk management, risk figures rely on single default probabilities and the dependence between the assets, usually modeled as correlation. The estimation of these parameters depends on the given data and with those certain implied risk figures. Parameter estimates change if data change, but how? Using the modeled distribution of the data we obtain asymptotic distributions for our parameter estimates describing the extend of these deviations. Taking these asymptotic distributions, we derive confidence intervals for our risk figures, namely probabilities of a percentage loss in a credit portfolio, the expected shortfall of Vasicek's loss distribution, the expected loss of a CDO-tranche etc..

1. Introduction

The need to advance a high-level modeling of credit markets has been strengthened in the last years as credit markets have become increasingly liquid. In particular, correlation products as Collateralized Debt Obligations (CDO) and structured credit derivatives definitely achieved to be a hot topic in mathematical finance. Apart from single default probabilities the price of correlation products is mainly driven by the dependence structure of the underlying portfolio. An appropriate model is required. It can be said that the Gaussian one factor model/normal copula model has become market standard in CDO pricing, although it does not perform very well. As it is assumed that the firm value is Gaussian, dependence reduces to an analysis of correlation. Estimating correlation is therefore of major importance in banks.

We are less interested in a price of a financial product. The aim of this paper is rather to examine properties of the frequently applied Gaussian copula model under the historical measure. In particular, we mainly focus on estimating default probabilities and correlations jointly, see also Bluhm and Overbeck (2003). We discuss classical estimation on historical loss rates

Date: July 7, 2006, revised: June 18, 2008.

We thank Sascha Meyer-Dautrich, Matthias Heck, Thomas Seifert, and Christoph Wagner for helpful comments and suggestions concerning this work. CN gratefully acknowledges financial support by UniCredit, Markets and Investment Banking. However, this paper does not reflect the opinion of UniCredit, Markets and Investment Banking or Commerzbank, it is the personal view of the authors.

E-mail address (corresponding author): Christina.Niethammer@uni-konstanz.de.

provided by S&P default study (2006) and derive certain risk figures, i.e. statistics of the considered portfolio loss distribution. For example we examine the expected loss of a CDO-tranche which would essentially drive the price of the tranche if the calibration were made to market prices, i.e. when expectations are taken under the risk-neutral measure. As our estimation is applied to historical loss rates, the implied risk figures are under the historical measure. From the obtained expected loss, a price therefore cannot be derived. However, a comparison to the market price gives a good intuition of the risk premium paid for the product in the market. Not least in view of the subprime crisis in 2007 this is of major interest to distinguish between actually bad credits and an indirect movement caused by adverse market conditions.

A major problem of such classical estimation techniques are the large estimation errors coming from a very small number of observations (here 25 periods from 1981-2005). Estimation errors in the parameter estimation are carried forward to estimation errors of the analyzed risk figures. This causes model risk. Reserves have to be set aside, additional costs are generated. We therefore build confidence intervals of the parameters implying confidence intervals for our risk figures to quantify the extend of potential model risk. This is established by deriving the (asymptotic) distribution of the parameters. In all cases, we can show that our parameters are asymptotically normal. (van der Vaart, 1998) serves as a good reference for the necessary background in mathematical statistics.

The paper is organized as follows. We start with some preliminaries on the considered Gaussian one factor model and the concept of identifying the effect on estimation errors on risk figures. We continue by comparing maximum likelihood and moment estimation in the Gaussian one factor model. Local asymptotic normality is proven by means of techniques presented in Le Cam and Yang (1990); van der Vaart (1998). Hence, we can examine asymptotic efficiency of the estimators, in the sense that the estimators can be asymptotically minimax, see Section 3.3. As we face a very small sample size, we further compare the estimators by means of a Monte Carlo simulation study. Moment estimation seems to perform better in finite samples. We therefore analyze the effect of estimation errors on risk figures by moment estimation only. A detailed empirical study is presented in the end. Some concluding remarks close the paper.

Remark 1.1 (Notation). *(i)* Y_t *denotes the common factor of the Gaussian one factor model in period* t. $(Y_t)_t$ *is assumed to be strictly stationary throughout the paper, hence* $E(f(Y_t))$ *is constant over time, provided* f *is designed such that the expectation of* $f(X_t)$ *exists. If it is clear from the context, we therefore often write* $E(f(Y))$ *instead of* $E(f(Y_t))$. f *will be in most cases equal to the identity or a function that maps into* $[0, 1]$ *so that we usually do not have to discuss the existence of the moments further.* $\mathbf{Y}$ *denotes the sequence* $(Y_t)_t$. *Again if it is clear from the context, we just* *write Y instead of $\mathbf{Y}$.*

(ii) $\mathcal{N}(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2. Its density is denoted by $f_{\mathcal{N}}(\cdot|\mu, \sigma^2)$. $X \sim \mathcal{N}(\mu(X), \sigma^2(X))$ means that X is normally distributed with expectation $\mu(X) = E(X)$ and variance $\sigma^2(X) := E(X^2) - (E(X))^2$.

2. Preliminaries

2.1. **Model.** We start with a short review on the Gaussian factor model in credit portfolio risk. We then specify the model further and explain the infinite granular assumption, i.e. we will assume that a large number of equal obligors is given. The object of interest will be therefore the loss rate within a specific rating class, observed in n different time periods. We assume that these loss rates are independent over time. Since the number of obligors is large, every sample of obligors within the considered periods is viewed as a good representation of the whole universe of obligors with a certain default probability specified by the rating class. We therefore consider every time period as an independent occurrence of the Gaussian one factor model (dependent on the default probability and correlation in the viewed rating class).

In a general factor model the risk driver $S_{i,t}$ (asset value return i at time t) is assumed to be dependent on factors $\check{\Phi}_{t,k},\ k = 1, ..., \check{G}$ and residual idiosyncratic risks $\epsilon_{t,i},\ i = 1, ..., N_t$ (N_t describes the number of obligors in period t):

$$S_{t,i} = \sum_k \check{w}_{i,k} \check{\Phi}_{t,k} + \epsilon_{t,i},\ t = 1, ..., n, \tag{1}$$

where $E(\epsilon_{t,j}\Phi_{t,k}) = 0,\ E(\epsilon_{s,i}\epsilon_{t,j}) = 0,\ t, s = 1, ..., n,\ i \neq j$. Factors can correspond to an underlying stochastic process or just a sequence of independent variables. In the latter case, the index t stands for the considered period. We focus on the estimation of a one factor model, we set $\check{G} = 1$, $S_{i,t} = X_{i,t},\ \check{\Phi}_{t,k} = Y_t,\ \check{w}_i = \sqrt{\rho}$, and $\sigma(\epsilon) = \sqrt{1-\rho}$ for all companies throughout this section. Y_t is assumed to be independent and standard normal. Extensions are left for future research. This Gaussian one factor model was introduced in finance by Vasicek (1991a). A single name perspective of the asset value model is given in Merton (1974). For further explanations see (Bluhm et al., 2002; Lando, 2004).

In detail, it is assumed that default at the end of period t occurs if a certain asset value process $X_{i,t}$ falls below a certain threshold $D_{i,t}$ (again at the end of period t). We assume that the asset value process $X_{i,t}$ for every firm is Gaussian at every time point and is independent of the other firms given the Gaussian distributed common factor Y_t. The joint common factor Y_t driving the asset value process is independent and identically distributed over time, hence all periods can be treated identically. In a follow-up paper we will allow for a certain dependent between Y_t and Y_s for $t \neq s$. To reduce the number of parameters, correlation structure is simplified extremely. The correlation between every asset i ($X_{i,t}$) and asset j ($X_{j,t}$) is set constant to $\rho \in (0,1)$. So on a one year horizon the asset (return) value process of firm i is described as:

$$X_{i,t} = \sqrt{\rho} Y_t + \sqrt{1-\rho} Z_{i,t}, \; i = 1,..,N_t \tag{2}$$

where $\rho \in (0,1)$, $Y_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $Z_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. (i.i.d. is an acronym for independent and identically distributed.) By now, we further assume that the one year default probability of company i at t, denoted as $p_{i,t}$, is known. Under this specification, we can easily derive the default threshold of company i, unconditionally we have with $H_{i,t} = 1_{\tau_i \leq t}$ and $D_{i,t}$ a deterministic number depending on the obligor and potentially on time:

$$\begin{aligned} P(\text{asset } i \text{ defaults at } t) &= P(H_{i,t} = 1) = P(X_{i,t} < D_{i,t}) = p_{i,t} \\ &\Rightarrow D_{i,t} = \Phi^{-1}(p_{i,t}) \end{aligned} \tag{3}$$

and given the common factor:

$$\begin{aligned} P(\text{asset } i \text{ defaults at } t|Y_t) &= P(X_{i,t} < D_{i,t}|Y_t) = p_{i|Y_t} \\ \Rightarrow g_{i,t}(Y_t) = p_{i|Y_t} = \Phi\left(\frac{\Phi^{-1}(p_{i,t}) - \sqrt{\rho} Y_t}{\sqrt{1-\rho}}\right) &\Rightarrow \int g_{i,t}(y)\phi(y)dy = p_{i,t} \end{aligned}$$

This factor model is equivalent to a credit model with Bernoulli random variables with *success* probabilities $p_{i,t}$ which are linked by a Gaussian copula with correlation ρ. So the model is also called one factor Gaussian copula model. Later, we replace $p_{i,t}$ by p_{R_k} (in short p_k) if company i is in rating class k in period t. Consequently, $g_{i,t}$ is then denoted by g_k. If there is only one rating class, we just write g. By the law of large number conditioned on the state of Y_t this motivates the infinite granular assumption for the loss rates in a portfolio of a large number of equal obligors within a rating class k:

Assumption 2.1 (Infinite Granular Assumption (IGA)). *The loss rate $\tilde{L}_{t,k}$ of an obligor in rating class k in period t possesses the form:*

$$\tilde{L}_{t,k} = g_k(Y_t) = \Phi\left(\frac{\Phi^{-1}(p_k) - \sqrt{\rho_k} Y_t}{\sqrt{1-\rho_k}}\right), \; Y_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1) \tag{4}$$

where $p_k := p_{R_k}$ is the one-period default probability in rating class R_k and ρ_k denotes the asset correlation within the obligors in rating class k for $k = 1, ..., K$.

Remark 2.1. *We wish to remark that in IGA the loss rate is modeled. Knowledge about the number $N_{t,k}$ is not necessary, i.e. there is no idiosyncratic component in this "asymptotic" approach. However, realizations of loss rates can be obtained dividing actual losses by the number of obligors in a rating class, i.e. $\tilde{L}_{t,k} = \frac{l_{t,k}}{N_{t,k}}$. $g_k(Y_t)$ is then the conditional default probability, but it is equal to the loss rate within the IGA when $N_{t,k}$ is "infinity".*

In the model (2) we already know that defaults of companies, given the common factor, are independent. We further restrict our consideration to the case when the losses of all firms are 1 and appear with probability $p = p_{i,t}$ for all i and t (the threshold D is equal for each firm). A generalization to rating classes is straightforward. The final one year default probability is then obtained by integrating out the common factor, i.e.

$$\forall t \forall i: \; P(H_{i,t} = 1) = \int_{-\infty}^{\infty} p_{|y} \phi(y) dy. \tag{5}$$

The default correlation $\tilde{r}$ between two assets is further given by:

$$\tilde{r} := \tilde{r}_{i,j,t} = Corr(H_{i,t}, H_{j,t}) = \frac{P(H_{i,t} = 1, H_{j,t} = 1) - p^2}{p(1-p)} \tag{6}$$

where $P(H_{i,t} = 1) = p$ and

$$P(H_{i,t} = 1, H_{j,t} = 1) = P(X_{i,t} < D_{i,t}, X_{j,t} < D_{j,t}) = \Phi_2(D_{i,t}, D_{j,t}, \rho) \tag{7}$$

with $D_{i,t} = \Phi^{-1}(p_{i,t}) := D$. Φ_2 denotes the standard bivariate normal distribution function. Moreover, the cumulated loss given the common factor is then a sum of i. i. d. Bernoulli random variables and therefore Binomial distributed (Bin), i.e.

$$\begin{aligned} L_{|Y_t} &= \sum_{i=1}^{N} 1_{X_{i,t} < D | Y_t} \\ P(L_t = l_t | Y_t) &= Bin(N_t, p_{|Y_t}, l_t). \end{aligned} \tag{8}$$

The final one year loss distribution is then obtained by integrating out the common factor, i.e.

$$P(L = l) = \int_{-\infty}^{\infty} Bin(N, p_{|y}, l) \phi(y) dy. \tag{9}$$

For a large number of assets this integral is computationally intensive, because of the binomial coefficient $\binom{N}{l}$. However, by the law of large numbers the cumulative distribution function of the percentage η portfolio loss $F_N(\eta)$ converges to the so called large pool loss distribution function $F_\infty(\eta)$, see Vasicek (1991b):

$$F_N(\eta) = \sum_{l=1}^{[N\eta]} P(L = l) \rightarrow F_\infty(\eta)$$

where

$$F_\infty(\eta) = \Phi\left(\frac{\sqrt{1-\rho}\Phi^{-1}(\eta) - D}{\sqrt{\rho}}\right) = P(g(Y) \leq \eta), \; Y \sim N(0,1). \tag{10}$$

Moreover, under the infinite granular assumption, our model is fully specified, and we obtain the following loss density, see also (Bluhm et al., 2002, p.91):

$$\tilde{L}_t = g(Y_t) = \Phi\left(\frac{\Phi^{-1}(p) - \sqrt{\rho}Y_t}{\sqrt{1-\rho}}\right),\ Y_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1) \tag{11}$$

$$\Rightarrow\quad \Phi^{-1}(\tilde{L}_t) \sim \mathcal{N}\left(\frac{\Phi^{-1}(p)}{\sqrt{1-\rho}}, \frac{\rho}{1-\rho}\right),\ \mu = \frac{\Phi^{-1}(p)}{\sqrt{1-\rho}},\ \sigma^2 = \frac{\rho}{1-\rho}$$

$$\Rightarrow\quad \tilde{L}_t \sim F_\infty = F_{\tilde{L},p,\rho}, F'_{\tilde{L},p,\rho} = f_{\tilde{L},p,\rho} := n_{\mu,\sigma^2}(\Phi^{-1}(\cdot))\frac{1}{\phi(\Phi^{-1}(\cdot))}$$

The explicit representation by a density will be very valuable when deriving the ML-estimator. Last but not least, the above one factor model can formally be included into a continuous setting, it corresponds to a dynamic underlying asset value process, given in Merton (1974) and Black and Scholes (1973) :

$$dV_{s,i}^{(t)} = \mu_i^{(t)} V_{s,i}^{(t)} ds + \sigma_i^{(t)} V_{s,i}^{(t)} dW_{s,i}, V_{t-1,i}^{(t)} = V^{(t)}(t-1,i), \tag{12}$$

for $i = 1, ..., N_t,\ s \in [t-1,t],\ t = 1, ..., n$ where W is an N_t-dimensional Brownian motion with pairwise correlation $\rho_{ij} = \rho$ for all $i \neq j$. The standardized log returns of $V^{(t)}$ in $[t-1,t]$ then correspond to $X_{i,t}$, i.e.

$$X_{i,t} = \frac{\log\left(\frac{V_t^{(t)}}{V_{t-1}^{(t)}}\right) - (\mu_{i,t} - \frac{1}{2}\sigma_{i,t}^2)}{\sigma_{i,t}}.$$

The correlation of $X_{i,t}$ and $X_{j,t}$ is ρ. (2) is just a reparametrization of $X_{i,t}$ by means of a one-factor-model, but only allowing for positive correlation.

Remark 2.2. *(i) In the definition of $X_{i,t}$ the starting point of V, i.e. $V^{(t)}(t-1,i)$ is factorized. This means the autocorrelation usually induced by geometric Brownian motion is not considered.*
(ii) In our statistical models we consider n periods with length 1 and regard them as n independent copies of an experiment. So in period $t = 1$, we consider N_1 companies with asset value processes as given in (12). In the next period $t = 2$, we examine N_2 firms with asset value processes given in (12), but potentially following another path of the above dynamics. Firms are allowed to be different from those in the first period, more specific the firm corresponding to an index i might be different in different time intervals $[t-1,t]$. Therefore one has to be careful in the dynamic interpretation of different versions or implementation of the asset value concept. Within IGA this makes perfectly sense as we assume that we have a sufficiently large number of obligors.

2.2. **The Effect on Estimation Errors - an Asymptotic Concept.** In this section, we sketch the main issue of this paper. We discuss implications of estimation errors on risk figures. Suppose we have already obtained estimators $(\hat{p}, \hat{\rho})$ in the Gaussian one factor model, e.g. moment or maximum likelihood estimation as discussed below. If the (asymptotic) distribution of the estimators is known, we can also obtain the (asymptotic) distribution of our risk figures provided that there are continuously differentiable functions of the true values p and ρ. This directly follows from the so called delta-method, see e.g. van der Vaart (1998, Theorem 3.8). Confidence intervals can be calculated. In our case both estimators of the considered parameters are even asymptotically normal as long as the true parameters do not lie on the boundary $\rho, p \neq 0/\rho, p \neq 1$.

We start with the definition of asymptotic normality. It is defined in the following sense:

Definition 2.1. *U_n is asymptotically normal with mean μ_n and covariance matrix Σ, denoted by $U_n \sim \mathcal{AN}(\mu_n, \delta_n^2\Sigma)$, if Σ is positive definite and there exists a sequence $\delta_n \to 0$ such that*

$$\frac{U_n - \mu_n}{\delta_n} \xrightarrow{d} \mathcal{N}(0, \Sigma).$$

If $\delta_n = \mathcal{O}(n^{-\beta})$, convergence is obtained with rate β. If $\delta_n = \mathcal{O}(e^{-\beta_e n})$, we say that convergence is obtained with exponential rate β_e.

We cite a special case of the delta method sufficient in our case, see also Serfling (1980, p.124):

Theorem 2.1. *Let Σ be a covariance matrix and $U_n = (U_{n1}, ..., U_{nk}) \sim \mathcal{AN}(\mu, \delta_n^2\Sigma)$, i.e. U_n is asymptotically normal with respect to the sequence $\delta_n \to 0$. Further suppose that $f(u)$ is a real-valued function having a nonzero differential at $u = \mu$. Then*

$$f(U_n) \sim \mathcal{AN}\left(f(\mu), \delta_n^2 \sum_{i=1}^{k}\sum_{j=1}^{k} \sigma_{ij} \left.\frac{\partial f}{\partial u_i}\right|_{u=\mu} \left.\frac{\partial f}{\partial u_j}\right|_{u=\mu}\right). \tag{13}$$

Suppose we consider a risk figure that can be described by a continuously differentiable function f of the true parameters p and ρ. By applying Theorem 2.1 and plugging in the estimators $\hat{p}$ and $\hat{\rho}$, we get an asymptotic confidence interval with level α for the considered risk figure $f(p, \rho)$:

$$\left(f(\hat{p}, \hat{\rho}) - \sigma_n\delta_n\Phi^{-1}(1 - \alpha/2), f(\hat{p}, \hat{\rho}) + \sigma_n\delta_n\Phi^{-1}(1 - \alpha/2)\right), \tag{14}$$

where $\sigma_n^2 = \sigma_{11}^2(\frac{\partial f}{dp})^2 + \sigma_{22}(\frac{\partial f}{d\rho})^2 + 2\sigma_{12}\frac{\partial f}{d\rho}\frac{\partial f}{dp}$ and provided that the vector $(\hat{p}, \hat{\rho})$ is asymptotically normal with δ_n.

3. Estimation in an Infinite Granular Model

In the remaining sections of this paper, we will restrict our considerations to the case when Assumption 2.1 holds, i.e. for the loss rate $\tilde{L}_{t,k}$ we have from (10) and (11):

$$\tilde{L}_{t,k} \sim F_\infty = F_{\tilde{L},p_k,\rho_k},\ f_{\tilde{L},p_k,\rho_k}(\cdot) = n_{\mu_k,\sigma_k^2}(\Phi^{-1}(\cdot))\frac{1}{\phi(\Phi^{-1}(\cdot))}$$

where $p_k := p_{R_k}$ is the one-period default probability in rating class R_k, $k = 1, ..., K$. Motivated by the Binomial model and assuming a large number of obligors, we have specified the distribution of the loss rate in dependence of the probability of default p_k and the asset correlation ρ_k. The information how many obligors are in one rating class can be ignored under this additional assumption. The model is fully specified by the distribution of $\tilde{L}_{t,k}$. We consider a moment estimator introduced in Bluhm and Overbeck (2003) and a maximum likelihood estimator:

(1) A *maximum likelihood* $(\hat{p}_{ML}, \hat{\rho}_{ML})$ estimate can be calculated as the maximizer of:

$$\max_{p_k,\rho_k} f_{\tilde{L},p_k,\rho_k}(\tilde{L})$$

(2) Bluhm and Overbeck (2003) propose a *moment estimator* $(m_{\tilde{L}}, \hat{\rho}_{BO}))$:

$$\hat{p}_k := m_{\tilde{L}(k)} = \frac{1}{n}\sum_{t=1}^{n} \tilde{L}_{t,k},\ s^2_{\tilde{L}(k)} = \frac{1}{n-1}\sum_{t=1}^{n}\left(\tilde{L}_{t,k} - \hat{p}_k\right)^2.$$

This moment estimator is robust to several model variants. To calculate an estimator of ρ, we will however need the infinite granular assumption. $\hat{\rho}_{BO}(k)$ is described as a function of $(\hat{p}_k, s^2_{\tilde{L}(k)})$ for $k = 1, ..., K$, see (21) below.

Both methods describe how to estimate default probability and asset correlation jointly. Every approach coincides for every rating class, we just skip the index k and write $p_k = p$, $\tilde{L}_{t,k} = \tilde{L}_t$. We further show that all estimators and other implied interesting risk figures are asymptotically normal. We start with the moment estimator:

3.1. **Moment Estimator in an Infinite Granular Model.** Let n the number of observed periods and $\tilde{L}_t$ $(t = 1, ..., n)$ the loss rate in IGA (Assumption 2.1), i.e. $\tilde{L}_t$ is a function of Y_t:

$$g(Y_t) := p_{|Y_t} = \tilde{L}_t. \tag{15}$$

Since Y_t are i.i.d. and $g(Y_t)$ is bounded in the unit interval, we have by the law of large numbers that

$$\overline{g^n(\mathbf{Y})} := \frac{1}{n}\sum_{t=1}^{n} g(Y_t) \overset{a.s.}{\rightarrow} E(g(Y)) = p, \tag{16}$$

$$\frac{1}{n-1}\sum_{t=1}^{n}(g(Y_t) - \overline{g^n(\mathbf{Y})})^2 \overset{a.s.}{\rightarrow} V(g(Y)), \tag{17}$$

and in addition by the central limit theorem

$$\sqrt{n}(\frac{1}{n}\sum_{t=1}^{n}\tilde{L}_t - p) \overset{d}{\rightarrow} \mathcal{N}(0, V(g(Y))) \tag{18}$$

and

$$\frac{1}{n}\sum_{s=1}^{n}\tilde{L}_s^2 \sim A\mathcal{N}(E(g^2(Y)), n^{-1}V(g^2(Y))).$$

As we consider i.i.d. random variables, we have joint asymptotic normality with covariance matrix $\sigma_{ij} = E(\tilde{L}^{i+j}) - E(\tilde{L}^j)E(\tilde{L}^i)$, see e.g. (Serfling, 1980, Theorem 2.2.1B). We follow Bluhm and Overbeck (2003) and define

$$m_{\tilde{L}} := \frac{1}{n}\sum_{t=1}^{n} \tilde{L}_t,\ s_{\tilde{L}}^2 := \frac{1}{n-1}\sum_{t=1}^{n}(\tilde{L}_t - m_{\tilde{L}})^2 \tag{19}$$

as reasonable estimates for $E(g(Y))$ and $V(g(Y))$. The estimator $\hat{p} := m_{\tilde{L}}$ of p is therefore asymptotically normal and consistent. Moreover, we have (in particular see (6))

$$\Phi_2(\Phi^{-1}(p), \Phi^{-1}(p), \rho) - p^2 = V(g(Y)). \tag{20}$$

Hence according to our model assumptions the true parameter ρ has to be the unique solution of (20). It uniquely exists for all $p \in (0,1)$ and $V(g(Y)) > 0$, see Lemma 3.1 below. To estimate ρ, p and $V(g(Y))$ are replaced by its sample moments $m_{\tilde{L}}$ and $s_{\tilde{L}}^2$. The estimate is then a function h of $m_{\tilde{L}}$ and $s_{\tilde{L}}^2$:

$$\hat{\rho}_{BO} := h_\rho(m_{\tilde{L}}, s_{\tilde{L}}^2), \tag{21}$$

where we replace h_ρ by h if its clear from the context. We are not able to determine h explicitly, but for our analysis (to derive asymptotic normality), we just need to know that h is continuously differentiable and calculate its first derivative at $p = E(g(Y))$ and $s = V(g(Y))$. This is established by the implicit functions theorem. Consequently, as h is continuous and $m_{\tilde{L}}$ and $s_{\tilde{L}}^2$ are consistent, $\hat{\rho}_{BO} = h(m_{\tilde{L}}, s_{\tilde{L}}^2)$ is consistent as well. Asymptotic normality of $\hat{\rho}_{BO}$ is shown next:

3.1.1. *Asymptotic Normality of* $\hat{\rho}_{BO}$. We show asymptotic normality of $\hat{\rho}_{BO}$ depending on $m_{\tilde{L}}$ and $s_{\tilde{L}}^2$. To use Theorem 2.1 we need to determine the first derivatives of h defined in (21).

To do so, we next apply the implicit function theorem to derive the derivatives of h, depending on h evaluated at (p,s). Suppose $\rho \in (0,1)$. We define

$$f : (0,1) \times (0,\infty) \times (0,1) \to \mathbb{R},\ f(p,s,\rho) = \Phi_2\left(\Phi^{-1}(p), \Phi^{-1}(p), \rho\right) - p^2 - s. \tag{22}$$

$(0,1) \times (0,\infty) \times (0,1)$ is open in itself. Further, we know that f is continuously differentiable infinitely often, i.e. $f \in \mathcal{C}^q,\ q \in \mathbb{N}$. Finally if

$$\frac{\partial f}{\partial \rho}(E(g(Y)), V(g(Y)), h(E(g(Y)), V(g(Y)))) \neq 0, \tag{23}$$

$$f(E(g(Y)), V(g(Y)), h(E(g(Y)), V(g(Y)))) = 0, \tag{24}$$

then there exists a neighborhood $U \subset (0,1)\times(0,\infty)$ and an $h \in \mathcal{C}^q(U,(0,1)), q \in \mathbb{N}$ such that

$$\forall u = (p,s) \in U : f(u,h(u)) = 0,\ \nabla h(u) = -\left(\frac{\partial f}{\partial \rho}\right)^{-1} \begin{pmatrix} \frac{\partial f}{\partial p} \\ \frac{\partial f}{\partial s} \end{pmatrix}(u).$$

To show (23) and (24), we start to derive $\frac{\partial f}{\partial \rho}$. By Vasicek (1996/98),

$$\frac{\partial}{\partial \rho}\Phi_2(z_1,z_2,\rho) = \phi_2(z_1,z_2,\rho) = (2\pi)^{-1}(1-\rho^2)^{-1/2}e^{-\frac{1}{2}\frac{z_1^2-2\rho z_1 z_2+z_2^2}{1-\rho^2}}.$$

Hence,

$$\frac{\partial f}{\partial \rho}(p,s,\rho) \quad = \quad \phi_2(\Phi^{-1}(p),\Phi^{-1}(p),\rho) = (2\pi)^{-1}(1-\rho^2)^{-1/2}e^{-\frac{(\Phi^{-1}(p))^2}{1+\rho}} > 0.$$

The solution of (20) therefore uniquely exists by (25) as $f(p,s,0) = -s$ and $f(E(g(Y)),V(g(Y)),1) = E(g(Y)-g^2(Y)) \geq 0$ for all $E(g(Y)) =: p \in (0,1)$, because $0 \leq g(Y) \leq 1$. This proves:

Lemma 3.1. *For $p = E(g(Y)) \in (0,1)$ with $V(g(Y)) > 0$ the following equation has a unique solution $\rho \in (0,1)$:*

$$\Phi_2(\Phi^{-1}(p),\Phi^{-1}(p),\rho) - p^2 = V(g(Y)).$$

In particular, Equation 23 and 24 hold.

Further, we have $\frac{\partial f}{\partial s} = -1$. It remains to derive $\frac{\partial f}{\partial p}$. As $\frac{\partial \Phi^{-1}(p)}{\partial p} = \frac{1}{\phi(\Phi^{-1}(p))}$, $\frac{\partial \tilde{f}}{\partial v} = \int_{-\infty}^{w}\phi_2(z_1,v,\rho)dz_1$, and $\frac{\partial \tilde{f}}{\partial w} = \int_{-\infty}^{v}\phi_2(w,z_2,\rho)dz_2$, where $\tilde{f}(w,v) = \Phi_2(w,v,\rho)$, by the chain rule we get:

$$\frac{\partial f}{\partial p}(p,s,\rho) \quad = \quad \frac{2}{\phi(\Phi^{-1}(p))}\int_{-\infty}^{\Phi^{-1}(p)}\phi_2(z_1,\Phi^{-1}(p),\rho)dz_1 - 2p \qquad (25)$$

$$\quad = \quad 2\mathcal{N}_{\rho\Phi^{-1}(p),1-\rho^2}(\Phi^{-1}(p)) - 2p \qquad (26)$$

where $\mathcal{N}_{\rho\Phi^{-1}(p),1-\rho^2}$ is the cumulative normal distribution with variance $1-\rho^2$ and mean $\rho\Phi^{-1}(p)$. Finally,

$$\nabla h(p,s) \quad = \quad -\left(\frac{e^{-\frac{(\Phi^{-1}(p))^2}{1+\rho}}}{2\pi\sqrt{1-\rho^2}}\right)^{-1}\begin{pmatrix} 2\mathcal{N}_{\rho\Phi^{-1}(p),1-\rho^2}(\Phi^{-1}(p)) - 2p \\ -1 \end{pmatrix}$$

where $\rho = h(p,s)$.

To show asymptotic normality of $\hat{\rho}_{BO}$, according to Theorem 2.1 we need joint asymptotic normality of $s^2_{\tilde{L}}$ and $m_{\tilde{L}}$. This is shown as follows: the joint asymptotic normality of $m_{\tilde{L}}$ and $\frac{1}{n}\sum_t \tilde{L}_t^2$ holds by (Serfling, 1980, Theorem 2.2.1B) as shown above. By multiplying $\frac{n}{n-1}$, we obtain the same result for $\frac{1}{n-1}\sum_{s=1}^n \tilde{L}_s^2$ as $\frac{n}{n-1}$ tends to 1. We then define $\tilde{h}(z_1,z_2) = z_1 - z_2^2, z_1 = \frac{1}{n}\sum_{s=1}^n \tilde{L}_s^2, z_2 = m_{\tilde{L}}$. Hence, by Theorem 2.1 and because $\frac{1}{n}\sum_{s=1}^n \tilde{L}_s^2$ and $\frac{1}{n-1}\sum_{s=1}^n \tilde{L}_s^2$ are asymptotically equal, we get

$$s^2_{\tilde{L}} \sim A\mathcal{N}(V(g(Y)), n^{-1}\sigma_S^2) \qquad (27)$$

where

$$\sigma_S^2 \quad = \quad V(g(Y))4E^2(g(Y)) + V(g^2(Y))\cdot 1 \qquad (28)$$

$$-4(E(g(Y)^3) - E(g(Y)^2)E(g(Y)))E(g(Y)). \qquad (29)$$

We further calculate the covariance between $m_{\tilde{L}}$ and $s^2_{\tilde{L}}$, again since $g(Y_t)$ are i.i.d.:

$$
\begin{aligned}
& Cov\left(m_{\tilde{L}}, \frac{n-1}{n} s^2_{\tilde{L}}\right) = Cov\left(m_{\tilde{L}}, \frac{1}{n}\sum_t \tilde{L}_t^2\right) - Cov\left(m_{\tilde{L}}, m^2_{\tilde{L}}\right) \\
= \; & \frac{1}{n^2}\sum_s Cov(\tilde{L}_s^2, \tilde{L}_s) - \frac{1}{n^3}\sum_s\sum_t\sum_r Cov(\tilde{L}_s, \tilde{L}_t\tilde{L}_r) \\
= \; & \frac{1}{n^2}\sum_s Cov(\tilde{L}_s^2, \tilde{L}_s) - \frac{2}{n^3}\sum_s\sum_{t\neq s} Cov(\tilde{L}_s, \tilde{L}_s\tilde{L}_t) - \frac{1}{n^3}\sum_s Cov(\tilde{L}_s, \tilde{L}_s^2) - 0 \\
= \; & \frac{1}{n}(E(g^3(Y)) - pE(g^2(Y))) - \frac{2}{n^3}\sum_s\sum_{t\neq s}(E(g^2(Y))p - p^3) + \mathcal{O}(n^{-2}) \\
= \; & \frac{1}{n}(E(g^3(Y)) - pE(g^2(Y))) - \frac{pV(g(Y))}{n} + \mathcal{O}(n^{-2})
\end{aligned}
$$

Finally, note that $f(m_{\tilde{L}}, \frac{1}{n}\sum_{s=1}^n \tilde{L}_s^2 - m^2_{\tilde{L}}, 1) = m_{\tilde{L}} > 0$. So using the last arguments and Lemma 3.1, we can determine the asymptotic distribution of $\hat{\rho}_{BO}$:

Theorem 3.2. *For $p := E(g(Y)) \in (0,1)$ with $V(g(Y)) > 0$ we have*

$$\hat{\rho}_{BO} = h(m_{\tilde{L}}, s^2_{\tilde{L}}) \sim A\mathcal{N}(h(p, V(g(Y))), n^{-1}\sigma^2_\rho) \tag{30}$$

where

$$
\begin{aligned}
\sigma^2_\rho \; = \; & V(g(Y))\left(\frac{\partial h}{\partial p}(p, V(g(Y)))\right)^2 + \sigma^2_S\left(\frac{\partial h}{\partial s}(p, V(g(Y)))\right)^2 && (31) \\
& + 2\sigma_{ms}\frac{\partial h}{\partial p}(p, V(g(Y)))\frac{\partial h}{\partial s}(p, V(g(Y))) && (32)
\end{aligned}
$$

and

$$\sigma_{ms} = (E(g^3(Y)) - pE(g^2(Y))) - pV(g(Y)). \tag{33}$$

In practical application, we replace $h(p, V(g(Y)))$ by $h(m_{\tilde{L}}, s^2_{\tilde{L}})$. Again justified by the law of large numbers, $E(g^j(Y))$ is replaced by $\frac{1}{n}\sum_t \tilde{L}_t^j$, $j = 2, 3$. As $m_{\tilde{L}}$ and $\hat{\rho}_{BO}$ are asymptotically normal the distribution of interesting risk figures can be derived, see Section 2.2 and 3.4.

We continue with the maximum likelihood estimator:

3.2. Maximum Likelihood Estimate in an Infinite Granular Model.

We keep on assuming a large number of obligors in our portfolio (infinite granular, 4), but discuss the ML-estimator instead of the moment estimator. Recall, $\tilde{L}_{t,k} \sim F_{\tilde{L},p_k,\rho_k}$ where $f_{\tilde{L},p_k,\rho_k} = n_{\mu_k,\sigma_k^2}(\Phi^{-1}(\cdot))\frac{1}{\phi(\Phi^{-1}(\cdot))}$, $\mu_k = \frac{\Phi^{-1}(p_k)}{\sqrt{1-\rho_k}}$, $\sigma_k^2 = \frac{\rho_k}{1-\rho_k}$, i.e. we consider the following log likelihood function:

$$\mathfrak{l}_a(p_k, \rho_k|\tilde{L}) = \log f_{\tilde{L},p_k,\rho_k}(\tilde{L}).$$

We can derive the ML-estimator for p_k and ρ_k. In the sequel, we again skip the index k and write $p := p_k$, $\tilde{L}_{t,k} = \tilde{L}_t$. We like to know the exact form of the "asymptotic" ML-estimator:

$$\begin{aligned} \frac{\partial \mathfrak{l}_a}{\partial p} &= \frac{\sqrt{1-\rho}}{\rho\phi(\Phi^{-1}(p))} \sum \left(\Phi^{-1}\left(\tilde{L}_t\right) - \frac{\Phi^{-1}(p)}{\sqrt{1-\rho}} \right) \stackrel{!}{=} 0 \\ &\Rightarrow \hat{p}^{\rho}_{ML} = \Phi\left(\frac{\sqrt{1-\rho}}{n} \sum_t \Phi^{-1}\left(\tilde{L}_t\right) \right) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial \mathfrak{l}_a}{\partial \rho} &= \frac{1}{2} \sum \left(\frac{\left(\Phi^{-1}\left(\tilde{L}_t\right) - \frac{\Phi^{-1}(p)}{\sqrt{1-\rho}}\right)^2}{\rho^2} + \frac{\Phi^{-1}(p)\left(\Phi^{-1}\left(\tilde{L}_t\right) - \frac{\Phi^{-1}(p)}{\sqrt{1-\rho}}\right)}{\rho\sqrt{1-\rho}} \right) \\ &\quad - \frac{n}{2}\left(\frac{1}{1-\rho} + \frac{1}{\rho} \right). \end{aligned} \tag{34}$$

The last term in the sum of (34) is zero if $\frac{\partial \mathfrak{l}_a}{\partial p} = 0$. Thus plugging $\hat{p}^{\hat{\rho}_{ML}}_{ML}$ in (34) yields that $\frac{\partial \mathfrak{l}_a}{\partial \rho}(\hat{p}^{\hat{\rho}_{ML}}_{ML}, \hat{\rho}_{ML}) = 0$ is equivalent to:

$$\begin{aligned} -\hat{\rho}^2_{ML} + \frac{1}{n} \sum \left(\Phi^{-1}\left(\tilde{L}_t\right) - \frac{\Phi^{-1}(\hat{p}_{ML})}{\sqrt{1-\hat{\rho}_{ML}}} \right)^2 (1-\hat{\rho}_{ML}) - \hat{\rho}_{ML}(1-\hat{\rho}_{ML}) = 0 \\ \Leftrightarrow \frac{1}{n} \sum \left(\Phi^{-1}\left(\tilde{L}_t\right) - \frac{1}{n} \sum_t \Phi^{-1}\left(\tilde{L}_t\right) \right)^2 (1-\hat{\rho}_{ML}) - \hat{\rho}_{ML} = 0 \end{aligned}$$

which is further equivalent to

$$\hat{\rho}_{ML} = \frac{\frac{1}{n}\sum\left(\Phi^{-1}\left(\tilde{L}_t\right) - \frac{1}{n}\sum_t \Phi^{-1}\left(\tilde{L}_t\right)\right)^2}{1 + \frac{1}{n}\sum\left(\Phi^{-1}\left(\tilde{L}_t\right) - \frac{1}{n}\sum_t \Phi^{-1}\left(\tilde{L}_t\right)\right)^2}.$$

In short, with $\hat{p}_{ML} := \hat{p}^{\hat{\rho}_{ML}}_{ML}$

$$\hat{p}_{ML} = \Phi\left(\frac{\overline{\Lambda}}{\sqrt{1 + \overline{\Lambda}_2 - \overline{\Lambda}^2}} \right), \quad \hat{\rho}_{ML} = \frac{\overline{\Lambda}_2 - \overline{\Lambda}^2}{1 + \overline{\Lambda}_2 - \overline{\Lambda}^2}$$

where

$$\overline{\Lambda} = \frac{1}{n} \sum_t \Phi^{-1}\left(\tilde{L}_t\right), \; \overline{\Lambda}_2 = \frac{1}{n} \sum_t \left(\Phi^{-1}\left(\tilde{L}_t\right)\right)^2.$$

Remark 3.1. *In practice, we do not have an infinite number of obligors: to calculate the estimator, we thus plug $\frac{l_t}{N_t}$ into $\tilde{L}_t$. If l_t is zero, we set the loss rate equal to one basis point (bp), i.e. a loss rate of 0.01%.*

We next like to prove that the ML-estimator is actually asymptotically normal. This is done by establishing typical regularity conditions as in (van der Vaart, 1998, Theorem 5.41). We essentially face Gaussian distributed random variables. $\mathfrak{l}_a$ is three times continuously differentiable with first derivative $\dot{\mathfrak{l}}_a$. Further, $\Phi^{-1}(\tilde{L}_t)$ is Gaussian and therefore possesses all moments. So by setting the parameter space equal to $\Theta = (0,1) \times (0,1)$ we always find for all $\theta_0 \in \Theta$ a sphere such that the third derivative of $\mathfrak{l}_a$ at $\theta_0 = (p_0, \rho_0)$ is dominated by an integrable function in this sphere around θ_0. The Fisher information $I_\theta = E(\dot{\mathfrak{l}}'_a \dot{\mathfrak{l}}_a)$ is well defined and invertible, see below. So finally, all regularity conditions of the ML-estimator are satisfied.

It remains to show that the Fisher information I_θ is invertible, i.e. $\det(I_\theta) > 0$. As $E\sum_{t=1}^{n}(\Phi^{-1}(\tilde{L}_t) - \frac{\Phi^{-1}(p)}{\sqrt{1-\rho}}) = 0$ and $E\left((\frac{\partial \mathfrak{l}_a}{\partial \rho})^2\right) = E\left(-\frac{\partial^2 \mathfrak{l}_a}{\partial \rho^2}\right)$ we get:

$$
\begin{aligned}
& \frac{\phi^2(\Phi^{-1}(p))}{n}\left(E\left(\left(\frac{\partial \mathfrak{l}_a}{\partial p}\right)^2\right)E\left(-\frac{\partial^2 \mathfrak{l}_a}{\partial \rho^2}\right) - E\left(\frac{\partial \mathfrak{l}_a}{\partial p}\frac{\partial \mathfrak{l}_a}{\partial \rho}\right)^2\right) \\
= \; & \frac{-1}{\rho}\left(\frac{-\rho}{1-\rho}\left(\frac{1-\rho}{\rho^3} + \frac{1}{\rho^2}\right) - \frac{1}{2(1-\rho)^2} + \frac{1}{2\rho^2} + \frac{-\Phi^{-1}(p))^2}{4\rho(1-\rho)^2}\right) - \frac{\Phi^{-1}(p))^2}{4\rho^2(1-\rho)^2} \\
= \; & \frac{1}{2(1-\rho)^2\rho^3} > 0
\end{aligned}
$$

Hence, $\det(I_\theta) = \frac{n}{\phi^2(\Phi^{-1}(p))}\frac{1}{2(1-\rho)^2\rho^3} > 0$. Finally, we have established that $(\hat{p}_{ML}, \hat{\rho}_{ML})$ is asymptotically normal with variance equal to the inverse of the Fisher-information. Asymptotic confidence intervals for our risk figures can be derived as described in Section 2.2. In fact, the distribution of $\overline{\Lambda}$ and $\overline{\Lambda}_2$ is known as $\Phi^{-1}(\tilde{L}_t)$ is normal. By applying the transformation theorem, an explicit density of the parameters and explicit confidence intervals can be derived. We refrain from such a lot calculus as the next section shows that moment estimation performs better in finite samples.

3.3. **Comparison: Efficiency vs. Empirical Evidence.** In this section, we compare the estimators so far described in Section 3. We start to examine its variation and bias. The less variation is determined the more efficient an estimator is supposed to be, provided all estimators possess the expectation of the parameter (otherwise a bias is included in the notion of efficiency). In particular, we like to know more about a potential efficiency of the moment estimator (BO) in comparison to the ML-estimator (ML). This is again done under an asymptotic perspective. Apart from this theoretical point of view, we perform a simulation study to see which parameters are actually recovered best in finite samples.

3.3.1. *A Short Note on Asymptotic Efficiency.* For an introduction to asymptotic efficiency and local asymptotic normality, we refer to Le Cam and Yang (1990) or van der Vaart (1998). As already noted in the last section, $\log f_{\tilde{L},p,\rho}(\tilde{L})$ is clearly continuously differentiable. The Fisher information I_θ is continuous in (p, ρ). We have differentiability in quadratic mean and therefore local asymptotic normality. We can examine asymptotic efficiency.

We have already proven regularity conditions given in (van der Vaart, 1998, Theorem 5.41), see Section 3.2. The expansion in (van der Vaart, 1998, Lemma 8.14) is therefore satisfied. So finally, all regularity conditions of the ML-estimator (ML) are fulfilled. The ML-estimator is best regular and therefore asymptotically efficient.

In fact, the ML estimator is more efficient than the proposed moment estimator in Bluhm and Overbeck (2003). To give a simple example we restrict our considerations to the parameter p. We need to show that $V(g(y))/n$ is strictly greater than $E(\frac{\partial^2 \mathfrak{l}_a}{\partial p^2})$:

As $\frac{\partial}{\partial p} \log \frac{1}{\phi(\Phi^{-1}(\tilde{L}_t))} = 0$, we get with $\sigma^2 = \frac{\rho}{1-\rho}$

$$\frac{\partial \mathfrak{l}_a}{\partial p} = \frac{\sum_{t=1}^n (\Phi^{-1}(\tilde{L}_t) - \frac{\Phi^{-1}(p)}{\sqrt{1-\rho}})}{\sigma^2 \sqrt{1-\rho}\phi(\Phi^{-1}(p))}.$$

Hence,

$$E\left(\left(\frac{\partial \mathfrak{l}_a}{\partial p}\right)^2\right) = \frac{E\left(\sum_{t=1}^n \left(\Phi^{-1}(\tilde{L}_t) - \frac{\Phi^{-1}(p)}{\sqrt{1-\rho}}\right)/\sigma\right)^2}{\phi^2(\Phi^{-1}(p))\sigma^2(1-\rho)} = \frac{n}{\rho\phi^2(\Phi^{-1}(p))}$$

as Y_t is i.i.d. $\mathcal{N}(0,1)$. Thus we get an efficient variance of

$$\tilde{V}_n(p,\rho) = \frac{\rho\phi^2(\Phi^{-1}(p))}{n}.$$

As the ML-estimator is efficient, we already know that $\tilde{V}_n(p,\rho)$ is equal or smaller than $\frac{1}{n}V(g(Y))$. Values of p and ρ are easily found such that $\tilde{V}_n(p,\rho)$ is strictly smaller than $\frac{1}{n}V(g(Y))$, set e.g. $p = 0.01$ and $\rho = 0.1$, then

$$\tilde{V}_n(p,\rho) = 2.8413\ 10^{-6},\ \frac{1}{n}V(g(Y)) = 9.2966\ 10^{-5}.$$

Note, similar examples can be found for both parameters jointly. For simplicity the last case actually assumes that ρ is known.

3.3.2. *Empirical Comparison.* We next examine the moment and ML-estimator empirically. The moment estimator is indexed by BO. The maximum likelihood estimator is subscribed by ML. At first, we rely on S&P default study (2006) and then perform a simulation study to compare the estimators.

	$\hat{p}_{BO} = m_{\tilde{L}}$	$\hat{\rho}_{BO}$	$\hat{p}_{ML}$	$\hat{\rho}_{ML}$
CCC	0.2292	0.1638	0.2334	0.3352
B	0.0512	0.0763	0.0519	0.0692
BB	0.0117	0.1032	0.0134	0.1889
BBB	0.0027	0.0650	0.0033	0.2136
A	0.0004	0.0747	0.0004	0.0959

TABLE 1. S&P default study (2006), (1981-2005): Moment vs. ML-estimator in an infinite granular model

In Table 1, we see that the estimation of default probabilities comes quite close for all methods relative to the small size. All numbers are in the usually stated ranges of default probabilities for the corresponding rating buckets. Unfortunately, the obtained values of ρ strongly depend on the method. It remains to find out which method actually delivers the right results for ρ. Surprisingly, the following simulation study shows that moment estimation seems to work a lot better than asymptotically efficient ML-estimation, see Table 2/3/4. In the next section, we therefore discuss the moment estimator only. The major concern of the next lines is to indicate that moment estimation might be preferred. A full analysis of the study (also containing an analysis of implied risk figures) is given in Section 4.3, but focusing on moment estimation only.

In detail, we perform four simulation studies to examine which study actually delivers correct results. In all studies we set

$$\mathbf{p} = \hat{p}^{data} = (0.2292\ 0.0521\ 0.0117\ 0.0027\ 0.0004).$$

The common factor Y_t and the idiosyncratic components $Z_{i,t}$ are sampled independently from a standard normal distribution. The number of obligors is drawn from a Poisson distribution with intensity 1000 for every period and rating class.

In the first two studies, we simulate 25 periods, so $t = 1, ..., 25$. In the first study we further set ρ equal to 0.08 for every rating class. Afterwards we plug in the estimated ρ from S&P default study (2006) via moment estimation (note, plugging in the estimated ρ from another estimate delivers similar results), see Table 1. The overall uncertainty is naturally reduced if n increases. We therefore repeat the same simulation studies, but increase n to 1000. All results are reported in Table 2/3/4.

	p	$\hat{p}_{BO} = m_{\tilde{L}}$	$\hat{p}_{ML}$	ρ	$\hat{\rho}_{BO}$	$\hat{\rho}_{ML}$
CCC	0.2292	0.2065	0.2068	0.1683	0.1532	0.1413
B	0.0512	0.0450	0.0449	0.0763	0.0846	0.0769
BB	0.0117	0.0091	0.0099	0.1032	0.0849	0.1274
BBB	0.0027	0.0025	0.0029	0.0650	0.0744	0.1834
A	0.0004	0.0003	0.0004	0.0747	0.0700	0.0818

BO=moment estimation, ML= maximum likelihood estimation, p, ρ=true parameters

TABLE 2. Simulation study with $n = 25$, $\rho_k \neq \rho_j$, and $Y \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$: a comparison between moment and ML-estimation

In all studies maximum likelihood estimation seems to underestimate the correlation for portfolios only containing CCC names (CCC portfolios) and significantly overestimates correlations for BB,BBB, and A portfolios. Only the estimate for B names seems to be reasonable. Except for one rating class (mainly B) the moment estimator produces a better fit of ρ. Especially for

	p	$\hat{p}_{BO} = m_{\tilde{L}}$	$\hat{p}_{ML}$	ρ	$\hat{\rho}_{BO}$	$\hat{\rho}_{ML}$
CCC	0.2292	0.2282	0.2282	0.1683	0.1603	0.1577
B	0.0512	0.0514	0.0514	0.0763	0.0774	0.0786
BB	0.0117	0.0117	0.0123	0.1032	0.1082	0.1414
BBB	0.0027	0.0027	0.0031	0.0650	0.0669	0.1563
A	0.0004	0.0004	0.0004	0.0747	0.0963	0.0929

BO=moment estimation, ML= maximum likelihood estimation, p, ρ=true parameters

TABLE 3. Simulation study with $n = 1000$, $\rho_k \neq \rho_j$, and $Y \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$: a comparison between moment and ML-estimation

	p	$\hat{p}_{BO} = m_{\tilde{L}}$	$\hat{p}_{ML}$	ρ	$\hat{\rho}_{BO}$	$\hat{\rho}_{ML}$
CCC	0.2292	0.2633	0.2633	0.0800	0.0823	0.0756
B	0.0512	0.0652	0.0648	0.0800	0.0945	0.0792
BB	0.0117	0.0158	0.0157	0.0800	0.1003	0.1044
BBB	0.0027	0.0045	0.0047	0.0800	0.0777	0.1050
A	0.0004	0.0007	0.0007	0.0800	0.0756	0.1248

BO=moment estimation, ML= maximum likelihood estimation, p, ρ=true parameters

TABLE 4. Simulation study with $n = 25$, $\rho = 0.08$, and $Y \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$: a comparison between moment and ML-estimation

$n = 1000$, this effect is manifested. Even for such a high sample, parameters estimates come a lot closer for moment estimation than for maximum likelihood estimation.

Remark 3.2. *(i) The maximum likelihood estimate is calculated by setting zero loss rates to one basis point. If we do the same for the moment estimation, there are just minor changes. The default probabilities do not change up to the fourth decimal place except for A ratings, here we observe a change of 0.0001. Similar for ρ, there is no change for CCC, B, and BB ratings. For BBB, we observe a change from 0.0619 to 0.0650, but for A portfolios ρ is estimated to be 0.0461 instead of 0.0747! Although, the change of $s^2_{\tilde{L}}$ for the A portfolio is just $5.7 * 10^{-7}$!*
(ii) All results based on S&P's default study have to be interpreted with care as the sample size (number of time periods) is quite small! There are longer time series available (e.g. by Moody's see (Bluhm and Overbeck, 2003)), however it is not guaranteed that there is no break point in the overall data structure which would influence our i.i.d. assumption.

3.4. **Moment Estimator and the Effect on Risk Figures.** Our simulation study performed in Section 3.3 showed a clear evidence for moment estimation. Moreover, efficiency of the ML-estimator is only gained asymptotically. Having in mind the very small sample size and the still bad results of the ML-estimator for a sample of $n = 1000$, we exclusively treat moment estimation in the sequel. We derive the exact form of the confidence bounds of interesting risk figures by obtaining asymptotic normality from the delta method (see Theorem 2.1). We derive all risk figures as functions $\bar{h}$ of $m_{\tilde{L}}$ and its empirical variance and set $\hat{\rho}_{BO} = h(m_{\tilde{L}}, s^2_{\tilde{L}})$ as defined in (21). Hence, it remains to derive the derivatives of these functions with respect to p and ρ, adjust by $\frac{\partial h}{\partial p}$ and $\frac{\partial h}{\partial s}$, and show that the gradient of $\bar{h}$ is nonzero. We consider the following risk figures: Vasicek's loss distribution, $F_\infty(\eta)$, expected loss of CDO-trances with attachment point K_1 and detachment point K_2, EL_{K_1,K_2}, and default correlations, $\tilde{r}_{i,j}$. From (10), we know that $F_\infty(\eta)$ is a function of p and ρ. So our estimates of the probability of η% loss depends on $m_{\tilde{L}}$ and $h(m_{\tilde{L}}, s^2_{\tilde{L}})$. We define

$$\begin{aligned} F_{\infty,\hat{\rho}}(\eta) &= \check{h}_\eta(m_{\tilde{L}}, s^2_{\tilde{L}}) \\ \text{where } \check{h}_\eta(p,s) &= \Phi\left(\frac{\sqrt{1-h(p,s)}\Phi^{-1}(\eta) - \Phi^{-1}(p)}{\sqrt{h(p,s)}}\right). \end{aligned}$$

At a fixed point in time the expected loss of a CDO tranche with attachment point K_1% and detachment point K_2% (at one single payment date, see below) is:

$$EL_{K_1,K_2} = \frac{\int_{K_1}^1 (x-K_1)dF(x) - \int_{K_2}^1 (x-K_2)dF(x)}{K_2 - K_1} \tag{35}$$

where F is the distribution function of accumulated losses upto the fixed time.

Example 3.1. *For $K_2 = 1$ and $K_1 = F^{-1}(\alpha)$ (the α-quantile or Value at Risk at α of the distribution F), EL_{K_1,K_2} corresponds to "the conditional Value at Risk" - an expected shortfall measure.*

Under our assumptions, i.e. a normal copula model in a large portfolio (LHP), we have (see also Kalemanova et al. (2005)):

$$EL_{K_1,K_2} = \frac{\Phi_2(-\Phi^{-1}(K_1), D, -\sqrt{1-\rho}) - \Phi_2(-\Phi^{-1}(K_2), D, -\sqrt{1-\rho})}{K_2 - K_1}$$

where $D = \Phi^{-1}(p)$. We can again describe $\hat{EL}_{K_1,K_2}$ as a function of $m_{\tilde{L}}$ and $s^2_{\tilde{L}}$:

$$\hat{EL}_{K_1,K_2} = \bar{r}(m_{\tilde{L}}, s^2_{\tilde{L}}),$$

where

$$\bar{r}(p,s) = \frac{\Phi_2(-\Phi^{-1}(K_1), D, -\sqrt{1-\rho}) - \Phi_2(-\Phi^{-1}(K_2), D, -\sqrt{1-\rho})}{K_2 - K_1}$$

with $\rho = h(p,s)$ and $D = \Phi^{-1}(p)$. Finally, default correlations are considered:

$$\tilde{r} := \tilde{r}_{i,j,t} = Corr(H_{i,t}, H_{j,t}) = \frac{\Phi_2(\Phi^{-1}(p), \Phi^{-1}(p), \rho) - p^2}{p(1-p)}. \tag{36}$$

For all risk figures, we obtain asymptotic normality:

Theorem 3.3. *Suppose the assumptions of Lemma 3.1 hold then for* $\bar{h} = \check{h}_\eta, \bar{r}, \tilde{r}$ *we have*

$$\bar{h}_\eta(m_{\tilde{L}}, s^2_{\tilde{L}}) \sim A\mathcal{N}(\bar{h}_\eta(p, V(g(Y))), n^{-1}\sigma^2_{\bar{h}}) \tag{37}$$

where

$$\begin{aligned}\sigma^2_{\bar{h}} &= V(g(Y))\left(\frac{\partial \bar{h}}{\partial p}(p, V(g(Y)))\right)^2 + \sigma^2_S\left(\frac{\partial \bar{h}}{\partial s}(p, V(g(Y)))\right)^2 \\ &\quad +2\sigma_{ms}\frac{\partial \bar{h}}{\partial p}(p, V(g(Y)))\frac{\partial \bar{h}}{\partial s}(p, V(g(Y)))\end{aligned} \tag{38}$$

and σ^2_S, σ_{ms} *are given in (29) and (33).*

To justify the above result, we next calculate the above gradients to show that they are nonzero. We start with Vasicek's loss function in a large pool:

3.4.1. *Asymptotic Normality of Vasicek's Loss Distribution.* By direct calculation, one gets $\nabla\check{h}_\eta$:

$$\begin{aligned}\frac{\partial \check{h}_\eta}{\partial p}(p,s) = \quad & \phi\left(\frac{\sqrt{1-h(p,s)}\Phi^{-1}(\eta) - \Phi^{-1}(p)}{\sqrt{h(p,s)}}\right) \\ \times \quad & \Bigg(-\frac{1}{2}(h(p,s))^{-\frac{3}{2}}\frac{\partial h}{\partial p}(p,s)(\sqrt{1-h(p,s)}\Phi^{-1}(\eta) - \Phi^{-1}(p)) \\ & + \frac{\left(\frac{-1}{2}\frac{\partial h}{\partial p}(p,s)(1-h(p,s))^{-\frac{1}{2}}\Phi^{-1}(\eta) - \frac{1}{\phi(\Phi^{-1}(p))}\right)}{\sqrt{h(p,s)}}\Bigg)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial \check{h}_\eta}{\partial s}(p,s) = \quad & \phi\left(\frac{\sqrt{1-h(p,s)}\Phi^{-1}(\eta) - \Phi^{-1}(p)}{\sqrt{h(p,s)}}\right) \\ \times \quad & \Bigg(-\frac{1}{2}(h(p,s))^{-\frac{3}{2}}\frac{\partial h}{\partial s}(p,s)(\sqrt{1-h(p,s)}\Phi^{-1}(\eta) - \Phi^{-1}(p)) \\ & -(h(p,s))^{-\frac{1}{2}}\frac{1}{2}\frac{\partial h}{\partial s}(p,s)(1-h(p,s))^{-\frac{1}{2}}\Phi^{-1}(\eta)\Bigg).\end{aligned}$$

Finally, if $\frac{\partial \check{h}_\eta}{\partial s}(p,s) = 0$ then $\frac{\partial \check{h}_\eta}{\partial p}(p,s) \neq 0$, because $\frac{\partial h}{\partial s} \neq 0$. Hence, Theorem 3.3 holds for $\bar{h} = F_\eta$.

3.4.2. *Asymptotic Normality of the Expected Loss of a CDO-tranche.* Recall, $\hat{EL}_{K_1,K_2} = \bar{r}(m_{\tilde{L}}, s^2_{\tilde{L}})$, where

$$\bar{r}(p,s) = \frac{\Phi_2(-\Phi^{-1}(K_1), D, -\sqrt{1-\rho}) - \Phi_2(-\Phi^{-1}(K_2), D, -\sqrt{1-\rho})}{K_2 - K_1}$$

with $\rho = h(p,s)$ and $D = \Phi^{-1}(p)$. Again, we calculate the gradient of $\bar{r}$: with $\frac{\partial - \sqrt{1-h(s,p)}}{\partial p} = \frac{1}{2}(1-h(s,p))^{-1/2}\frac{\partial h}{\partial p}$, $\frac{\partial \Phi_2(w,v,\rho)}{\partial \rho} = \phi_2(w,v,\rho)$, $\frac{\partial \Phi_2(w,v,\rho)}{\partial v} = \int_{-\infty}^{w} \phi_2(u,v,\rho)du$ and the chain rule we get:

$$\begin{aligned} & \breve{r}(p,s,K) \\ := & \left(\frac{\partial \tilde{f}_w}{\partial v}(\Phi^{-1}(p), -\sqrt{1-\rho}) \;\; \frac{\partial \tilde{f}_w}{\partial \rho}(\Phi^{-1}(p), -\sqrt{1-\rho}) \right) \begin{pmatrix} \frac{\partial \Phi^{-1}(p)}{\partial p} \\ \frac{\partial -\sqrt{1-\rho}}{\partial p} \end{pmatrix} \\ = & \mathcal{N}_{-\sqrt{1-\rho}\Phi^{-1}(p),\rho}(-\Phi^{-1}(K)) \\ & +\phi_2(-\Phi^{-1}(K), \Phi^{-1}(p), -\sqrt{1-\rho})\frac{1}{2}(1-\rho)^{-1/2}\frac{\partial h}{\partial p} \end{aligned}$$

where $\rho = h(p,s)$ and $w = -\Phi^{-1}(K)$. Finally,

$$\frac{\partial \bar{r}}{\partial p}(p,s) \quad = \quad \frac{\breve{r}(p,s,K_1) - \breve{r}(p,s,K_2)}{K_2 - K_1}$$

and

$$\frac{\partial \bar{r}}{\partial s}(p,s) = \frac{\frac{\partial h}{\partial s}(\phi_2(-\Phi^{-1}(K_1), D, -\sqrt{1-\rho}) - \phi_2(-\Phi^{-1}(K_2), D, -\sqrt{1-\rho}))}{2\sqrt{(1-\rho)}(K_2 - K_1)}$$

where $D = \Phi^{-1}(p)$ and $\rho = h(p,s)$. Again if $\frac{\partial \bar{r}}{\partial s} = 0$ and $K_1 \neq K_2$ then $\frac{\partial \bar{r}}{\partial p} \neq 0$, as $\frac{\partial h}{\partial s} \neq 0$. Theorem 3.3 follows.

3.4.3. *Asymptotic Normality of Default Correlations.* Recall,

$$\tilde{r} := \tilde{r}_{i,j,t} = Corr(H_{i,t}, H_{j,t}) = \frac{\Phi_2(\Phi^{-1}(p), \Phi^{-1}(p), \rho) - p^2}{p(1-p)}. \tag{39}$$

Hence, $\tilde{r}$ is a function $\hat{h}$ of $m_{\tilde{L}}$ and $s^2_{\tilde{L}}$ (via ρ). From (26) (slightly modified, ρ is a function of p and s), we have

$$\begin{aligned} & \frac{\partial \hat{h}}{\partial p} \\ = & \frac{(2N_{\rho D, 1-\rho^2}(D) - 2p + \phi_2(D,D,\rho)\frac{\partial h}{\partial p})p(1-p) - (\Phi_2(D,D,\rho) - p^2)(1-2p)}{(p(1-p))^2} \end{aligned}$$

where $D = \Phi^{-1}(p)$. As $\frac{\partial h}{\partial s} \neq 0$ we have

$$\frac{\partial \hat{h}}{\partial s} = \frac{\phi_2(D,D,\rho)\frac{\partial h}{\partial s}}{p(1-p)} \neq 0. \tag{40}$$

4. Empirical Results

This section is devoted to an empirical study of S&P's default data. The empirical and the simulation study of Section 3.3 is resumed and extensively discussed for moment estimators. We estimate the default probability and the asset correlation for every rating class and calculate confidence intervals for our risk figures. A look at Table 2/3/4 shows that the approximation of a large number of obligors does not work really well applied to ML-estimation $(\hat{p}_{ML}, \hat{\rho}_{ML})$ in a simulation study. In contrary to the asymptotically efficient ML-estimatior, true values of our simulation study are recovered quite well by the moment estimator $(m_{\tilde{L}}, \hat{\rho}_{BO})$. In view of Section 3.3 we restrict our analysis on infinite granular and focus on moment estimators in this section only. We shortly say estimator.

For the reader's convenience, we again recall the method the following numbers are obtained from, see also Section 3.3.

4.1. **Notation and Calculation Method.** We start reviewing some notation. Our observations, provided by S&P default study (2006), are the numbers of companies in our rating classes $N_t := (N_{1,t}, ..., N_{K,t})$ and the numbers of defaults in these rating classes $l_t := (l_{1,t}, ..., l_{K,t})$ in period $t \in \{1, ..., n\}$, where $n = 25$ is the number of periods with $t = 1981 - 2005$ and $K = 5$ with $k = CCC, B, BB, BBB, A$. The corresponding loss random variable is denoted by $L_t = (L_{1,t}, ..., L_{K,t})$. Note, for moment estimation only the ratio of l and N is necessary, see also Remark 2.1. We assume that every company i in period t is assigned to a rating class $k_{i,t} \in \{R_1, .., R_K\}$. In rating class $R_k,\ k = 1, .., K$ a default appears with probability $p_k = p_{R_k}$. Parameters are thus the correlation parameters $\rho = (\rho_1, ..., \rho_K)$ and default probabilities of the different rating classes $\mathbf{p} = (p_1, ..., p_K)$.

For $k = 1, ..., K$ an estimator of $p_k = E(g_k(Y))$ is given by

$$m_{\tilde{L}(k)} = \frac{1}{n} \sum_{t=1}^{n} \frac{L_{t,k}}{N_{t,k}}.$$

A realization of $m_{\tilde{L}(k)}$, $m_{\tilde{L}(k)}(\omega)$, is equal to $\frac{1}{n} \sum_{t=1}^{n} \frac{l_{t,k}}{N_{t,k}}$. $V(g_k(Y))$ can be estimated by

$$s^2_{\tilde{L}(k)} = \frac{1}{n-1} \sum_{t=1}^{n} \left(\frac{L_{t,k}}{N_{t,k}} - m_{\tilde{L}(k)} \right)^2.$$

$\hat{\rho}_{BO}$ is then a function h of $m_{\tilde{L}(k)}$ and $s^2_{\tilde{L}(k)}$. The definition of h is given in (21). Note, all parameter estimates are calculated for every rating class separately. The estimators of rating class k contain no information about rating i and vice versa if $i \neq k$. So as shown in Section 3.1/3.4, we get estimates of ρ_k and asymptotic normality for the parameter estimates, the loss distribution, and the expected loss. We next list the obtained confidence intervals. The index k hereby denotes the k-th rating class. For p_k, we get

$$\left(m_{\tilde{L}}(k) - \frac{\sqrt{s^2_{\tilde{L}}(k)} \tilde{K}_{1-\alpha/2}}{\sqrt{n}}, m_{\tilde{L}}(k) + \frac{\sqrt{s^2_{\tilde{L}}(k)} \tilde{K}_{1-\alpha/2}}{\sqrt{n}} \right), \tag{41}$$

where $\tilde{K}_{1-\alpha/2} = \Phi^{-1}(1-\alpha/2)$. For ρ_k by Theorem 3.2, we obtain with $\hat{\rho}_{BO}(k) = h(m_{\tilde{L}(k)}, s^2_{\tilde{L}(k)})$:

$$\left(\hat{\rho}_{BO}(k) - \frac{\hat{\sigma}_{\rho,k}\tilde{K}_{1-\alpha/2}}{\sqrt{n}}, \hat{\rho}_{BO}(k) + \frac{\hat{\sigma}_{\rho,k}\tilde{K}_{1-\alpha/2}}{\sqrt{n}}\right) \quad (42)$$

where $\hat{\sigma}^2_{\rho,k}$ is generated by σ^2_S from Theorem 3.2 taking $m_{\tilde{L}}$ and $s^2_{\tilde{L}}$ as estimators of k-rated companies only.

For the loss distribution at η, we get with $F_{\infty,\hat{\rho}_{BO}(k)}(\eta) = \check{h}_\eta(m_{\tilde{L}(k)}, s^2_{\tilde{L}(k)})$ and $\sigma_{F(\eta),k}(m_{\tilde{L}(k)}, s^2_{\tilde{L}(k)}) = \hat{\sigma}_{F,\eta,k}$:

$$\left(F_{\infty,\hat{\rho}_{BO}(k)}(\eta) - \frac{\hat{\sigma}_{F,\eta,k}\tilde{K}_{1-\alpha/2}}{\sqrt{n}}, F_{\infty,\hat{\rho}_{BO}(k)}(\eta) + \frac{\hat{\sigma}_{F,\eta,k}\tilde{K}_{1-\alpha/2}}{\sqrt{n}}\right) \quad (43)$$

see Theorem 3.3.

Similar, for the expected loss of a CDO-Tranche with attachment point K_1 and detachment point K_2, we get with $EL_k = r_{K_1,K_2}(m_{\tilde{L}(k)}, s^2_{\tilde{L}(k)})$ and $\hat{\sigma}_{EL,k} = \sigma_{EL,K_1,K_2,k}(m_{\tilde{L}(k)}, s^2_{\tilde{L}(k)})$:

$$\left(EL_k - \frac{\hat{\sigma}_{EL,k}\tilde{K}_{1-\alpha/2}}{\sqrt{n}}, EL_k + \frac{\hat{\sigma}_{EL,k}\tilde{K}_{1-\alpha/2}}{\sqrt{n}}\right). \quad (44)$$

Finally note, we shortly say “CCC portfolio tranche” meaning a tranche of an underlying portfolio only containing CCC names.

4.2. **Application to S&P's Default Study.** In Table 5, we start with the estimation of p_k, $k = CCC, B, BB, BBB, A$ with empirical quantiles of $\hat{p}^{data} := m_{\tilde{L}}$. We further display $\sqrt{s^2_{\tilde{L}}(k)}$ which is an estimate of the standard deviation of $m_{\tilde{L}}(k) =: \hat{p}^{data}(k)$. We finally compare empirical quantiles with the asymptotic confidence intervals of $m_{\tilde{L}}(k)$. The first two columns show the empirical quantiles of $m_{\tilde{L}}(k)$. The asymptotic confidence bounds, a^p and b^p, are derived in Section 3.4. We see that the asymptotic confidence intervals are a lot tighter than the empirical ones, which can be explained by the rather small sample size of 25 periods.

p	$m_{\tilde{L}}$	$s^2_{\tilde{L}}$	$\sqrt{s^2_{\tilde{L}}}$	2.5%	median	97.5%	$a^p_{2.5\%}$	$b^p_{97.5\%}$
CCC	0.2292	0.0157	0.1254	0.0300	0.2143	0.4422	0.1800	0.2784
B	0.0512	0.0010	0.0309	0.0157	0.0409	0.1188	0.0400	0.0642
BB	0.0117	0.0001	0.0112	0	0.0095	0.0370	0.0073	0.0160
BBB	0.0027	8.2E-06	0.0029	0	0.0023	0.0087	0.0015	0.0038
A	0.0004	5.0E-07	0.0007	0	0	0.0019	0.0001	0.0007

TABLE 5. S&P default study: default probabilities

Table 6 reports an estimator of ρ_k for $k = CCC, B, BB, BBB, A$ with asymptotic confidence bounds. In our simulation study we will see that the true value sometimes does not lie in these asymptotic confidence bounds (usually for better ratings). The above observed effect that asymptotic confidence intervals are too tight is carried forward to the asymptotic confidence intervals of ρ for better rating classes. However, empirical deviation or quantiles are not available, because ρ_k is the solution of one equation.

ρ	$\hat{\rho}^{data} := \hat{\rho}^{data}_{BO}$	$a^{\rho}_{2.5\%}$	$b^{\rho}_{97.5\%}$
CCC	0.1638	0.0536	0.2740
B	0.0763	0.0407	0.1119
BB	0.1032	0.0897	0.1167
BBB	0.0650	0.0601	0.0700
A	0.0747	0.0731	0.0764

TABLE 6. S&P default study: estimation of asset correlations

We continue calculating confidence intervals of Vasicek's loss distribution and the expected loss of CDO-tranches in the same manner. At first, we present the estimated probability of η% loss in the portfolio, see Table 7. Next, asymptotic confidence intervals of the same points of Vasicek's loss distribution are determined ($\alpha = 0.05$, 2.5% on both sides). For η equal to 0.001 % this makes sense for A rated companies only. We obtain a confidence interval of $(a, b) = (0, 0.0101)$. For all others the right boundary b is smaller than 10^{-6}. Results are reported in Table 8 and 9 starting from $\eta = 0.01\%$. Naturally, if η is around $\hat{p} = m_{\tilde{L}}$ uncertainty increases.

$\check{h}_\eta/\eta\%$	0.001	0.01	0.1	2.5	5	10	25
CCC	$< 10^{-14}$	$< 10^{-10}$	0.13E-06	0.0047	0.0298	0.1438	0.6211
B	$< 10^{-14}$	$< 10^{-10}$	0.56E-06	0.1743	0.5632	0.9226	0.9998
BB	$< 10^{-07}$	0.47E-04	0.0202	0.9001	0.9865	0.9995	1.000
BBB	$< 10^{-07}$	0.0007	0.2137	0.9998	1.000	1.000	1.000
A	0.0032	0.2110	0.9207	1.000	1.000	1.000	1.000

TABLE 7. S&P default study: Vasicek's Loss Distribution $\check{h}_\eta$

A more interesting risk figure is the expected loss of a financial product. At first we consider the expected loss over one year of CDO-tranches. Recall, numbers are derived from S&P default study (2006), i.e. from historical loss rates. The obtained expected loss is therefore under the real world measure and does not directly drive the price of a CDO. It though gives an intuition about the risk premium included in the market prices. The equity tranche is not worth to look at for bad ratings. Similar for good ratings, the expected loss over a one year horizon in the 3-6% tranche is very small and not sensible to consider. We therefore map tranche attachement point

$\check{h}_\eta/\eta\%$	$a_{0.01}$	$b_{0.01}$	$a_{0.1}$	$b_{0.1}$	$a_{2.5}$	$b_{2.5}$
CCC	0	$< 10^{-9}$	0	0.14E-05	0	0.0179
B	0	$< 10^{-9}$	0	0.40E-05	0.0550	0.2937
BB	0	0.0001	0	0.0429	0.8203	0.9798
BBB	0	0.0021	0.0575	0.3698	0.9993	1.0000
A	0.0102	0.4119	0.8181	1.0000	1.0000	1.0000

TABLE 8. S&P default study: loss distribution (confidence bounds)

$\check{h}_\eta/\eta\%$	a_5	b_5	a_{10}	b_{10}	a_{25}	b_{25}
CCC	0	0.0803	0.0232	0.2644	0.4641	0.7780
B	0.4007	0.7256	0.8459	0.9994	0.9991	1.0000
BB	0.9702	1.0000	0.9986	1.0000	1.0000	1.0000
BBB	1.000	1.000	1.000	1.000	1.000	1.000
A	1.000	1.000	1.000	1.000	1.000	1.000

TABLE 9. S&P default study: loss distribution (confidence bounds)

to the credit quality of the underlying portfolio. For a portfolio with B rated companies considering a 3-6% tranche seems to be reasonable. The corresponding tranche is calibrated via a so called expected loss mapping to the same *moneyness* (here the expected loss) as the B portfolio tranche (i.e. $K_B = K_{CCC}EL(B)/EL(CCC) = K_{CCC}PD(B)/PD(CCC)$). A mapping of the other tranches makes no sense as detachment points would be very small. We therefore consider the equity tranche for the remaining rating classes. Results are reported in Table 10.

$EL^{K_2\%}_{K_1\%}$	K_1	K_2	$a^{EL}_{2.5\%}$	EL^{data}	$b^{EL}_{97.5\%}$
CCC	14	29	0.3335	0.4888	0.6441
B	3	6	0.3534	0.5156	0.6777
BB	0	3	0.2445	0.3608	0.4772
BBB	0	3	0.0514	0.0889	0.1264
A	0	3	0.0041	0.0131	0.0221

TABLE 10. S&P default study: expected loss of CDO tranches mapped to moneyness

So far we have considered periods of one year only. Clearly, on a one year horizon it is quite safe to invest in a portfolio of A rated companies. So more interesting is the probability that a company defaults within a longer time horizon. Itraxx tranches e.g. have been become firstly liquid on a five year horizon. However, if we want to have uncorrelated periods, just five of the 25 periods are left for an analysis. Alternatively, one can analyze five year default rates every year. In this case 21 periods are remaining. This overlap causes autocorrelation; therefore the numbers in Table 11 should be interpreted with care. Results are reported in the common notation of banks in basis points (bps). One basis point is equal to 0.01%. Expected loss is reported over the time period of five years, numbers are not annualized. We assume a recovery rate R of 0%, a generalization to 40% is straightforward by adjusting the (de)attachment points.

$EL_{K_1\%}^{K_2\%}$ in bps	$a_{2.5\%}^{EL_{3\%}^{6\%}}$	$EL_{3\%}^{6\%}$	$b_{97.5\%}^{EL_{3\%}^{6\%}}$	$a_{2.5\%}^{EL_{0\%}^{3\%}}$	$EL_{0\%}^{3\%}$	$b_{97.5\%}^{EL_{0\%}^{3\%}}$
BB	8620	9646	10000	9959	9993	10000
BBB	306	1180	2053	6571	7645	8720
A	1.2	3.47	5.7	1546	2109	2672
AA	0	2.87	8.5	502	1033	1563

TABLE 11. S&P default study: five year expected loss of 0-3% and 3-6% tranches in bps, zero recovery

4.3. **Simulation Study.** To guarantee that our method actually delivers correct results and to further analyze these results, we resume the four simulation studies from Section 3.3. Recall, in all studies we set

$$\mathbf{p} = \hat{p}^{data} = (0.2292,\ 0.0521\ 0.0117\ 0.0027\ 0.0004).$$

The common factor Y_t and the idiosyncratic components $Z_{i,t}$ are sampled independently from a standard normal distribution. The number of obligors is drawn from a Poisson distribution with intensity 1000 for every period and rating class.

The first two studies simulate 25 periods, so $t = 1, ..., 25$. In the first study we further set ρ equal to 0.08 for every rating class. Afterwards we plug in the estimated ρ from S&P default study (2006) in the last section. In Section 3.3, we have already seen that in both studies parameters are recovered quite well. In this section, we concentrate on the fluctuation of the parameter estimates, i.e. we examine asymptotic and empirical confidence intervals. See Table 12, Table 13, and Table 14 for the results of the simulation studies with $n = 25$. The overall uncertainty is naturally reduced if n increases. We therefore repeat the same simulation studies, but increase n to 1000. Results are reported in Table 15 and 16. The difference of the estimate $\hat{p}^{sim}$ to the input data ($\hat{p} = m_{\tilde{L}}$ for the data) is only recognizable for portfolios of CCC (0.001) and B (0.0007) rated companies. Finally, we analyze how good the estimate of the expected shortfall actually is.

4.4. **Discussion of the Empirical Results.** This section is devoted to a detailed examination of the results: we recognize different ρ for different rating classes and discuss implications. Further, for better ratings confidence intervals are not very accurate, we add possible reasons. In the next section, we then present some possible extensions to the implications arising from our study. Related literature is given.

p	$\hat{p}^{data} = m_{\tilde{L}}$	$\hat{p}^{sim}$	$s^2_{\tilde{L}}$	$\sqrt{s^2_{\tilde{L}}}$	2.5%	median	97.5%
CCC	0.2292	0.2633	0.0089	0.0943	0.1281	0.2247	0.4507
B	0.0512	0.0652	0.0017	0.0411	0.0241	0.0521	0.1543
BB	0.0117	0.0158	0.0002	0.0141	0.0028	0.0094	0.0460
BBB	0.0027	0.0045	0.0000	0.0042	0.0007	0.0030	0.0134
A	0.0004	0.0007	0.0000	0.0012	0	0	0.0034

TABLE 12. Simulation study with $n = 25$, $Y \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $\rho_k = 0.08$ for all ratings: estimated default probabilities

p	$\hat{p}^{data} = m_{\tilde{L}}$	$\hat{p}^{sim}$	$s^2_{\tilde{L}}$	$\sqrt{s^2_{\tilde{L}}}$	2.5%	median	97.5%
CCC	0.2292	0.2065	0.0131	0.1145	0.0550	0.1943	0.3956
B	0.0512	0.0450	0.0009	0.0293	0.0125	0.0362	0.1013
BB	0.0117	0.0091	0.0001	0.0080	0.0007	0.0067	0.0247
BBB	0.0027	0.0025	0.0000	0.0029	0	0.0010	0.0085
A	0.0004	0.0003	0.0000	0.0006	0	0	0.0013

TABLE 13. Simulation study with $n = 25$, $Y \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $\rho_k \neq \rho_j$: estimated default probabilities

	$\rho_{.08}$	$\hat{\rho}$	$a^{\rho}_{2.5\%}$	$b^{\rho}_{97.5\%}$	ρ_d	$\hat{\rho}$	$a^{\rho}_{2.5\%}$	$b^{\rho}_{97.5\%}$
CCC	0.0800	0.0823	0	0.1988	0.1683	0.1532	0.0524	0.2539
B	0.0800	0.0945	0.0518	0.1372	0.0763	0.0846	0.0521	0.1172
BB	0.0800	0.1003	0.0838	0.1167	0.1032	0.0849	0.0740	0.0958
BBB	0.0800	0.0777	0.0710	0.0844	0.0650	0.0744	0.0695	0.0793
A	0.0800	0.0756	0.0732	0.0781	0.0747	0.0700	0.0686	0.0715

TABLE 14. Simulation study with $n = 25$, $Y \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$ and different specifications of ρ.

Estimation of $\rho_{.08}$ within the variant with $\rho_k^{.08} = 0.08$ for all ratings is displayed in column 2 to 4. In column 6 to 9 one finds the estimation of ρ_d within the variant with $\rho_d := \hat{\rho}^{data}$ from Table 6.

In our simulation study all parameters are recovered for several studies of this type, we conclude that the procedure is able to find the true parameters. In particular, if ρ_k is set to the estimated $\hat{\rho}_{BO}(k)$ from the empirical study, $\hat{\rho}_{BO}(k)$ of the simulation study is quite close to ρ_k for every rating class. If $\rho_k \equiv 0.08$ for every rating class, it turns out that all estimates fluctuate around 0.08. On the other hand, in our empirical study ρ_k for different rating classes have been significantly different. Especially the correlation within the CCC portfolio is higher than for other ratings. Our simulation study

	$p = \hat{p}^{data}$	$\hat{p}^{sim}$	$s^2_{\tilde{L}}$	$\sqrt{s^2_{\tilde{L}}}$	2.5%	median	97.5%
CCC	0.2292	0.2282	0.0153	0.1238	0.0681	0.2063	0.4621
B	0.0521	0.0514	0.0010	0.0308	0.0153	0.0437	0.1108
BB	0.0117	0.0117	0.0001	0.0115	0.0010	0.0081	0.0341
BBB	0.0027	0.0027	0.0000	0.0030	0	0.0020	0.0089
A	0.0004	0.0004	0.0000	0.0008	0	0	0.0020

TABLE 15. Simulation study with $n = 1000$, $Y \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $\rho_k \neq \rho_j$: default probabilities

ρ	$\hat{\rho}^{data}$	$\hat{\rho}^{sim}$	$a^{\rho}_{2.5\%}$	$b^{\rho}_{97.5\%}$	ρ	$\hat{\rho}$	$a^{\rho}_{2.5\%}$	$b^{\rho}_{97.5\%}$
CCC	0.1638	0.1603	0.1434	0.1771	0.0800	0.0768	0.0605	0.0930
B	0.0763	0.0774	0.0719	0.0829	0.0800	0.0800	0.0744	0.0857
BB	0.1032	0.1082	0.1060	0.1103	0.0800	0.0798	0.0790	0.0806
BBB	0.0650	0.0669	0.0661	0.0677	0.0800	0.0798	0.0790	0.0806
A	0.0747	0.0963	0.0960	0.0965	0.0800	0.0936	0.0933	0.0939

TABLE 16. Simulation study with $n = 1000$, $Y \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and different specifications of ρ.

Estimation of $\rho_{.08}$ within the variant with $\rho_k^{.08} = 0.08$ for all ratings is displayed first in column 2 to 4. In column 6 to 9 one finds the estimation of ρ_d within the variant with $\rho_d := \hat{\rho}^{data}$ from Table 6.

$EL^{K_2\%}_{K_1\%}$	K_1	K_2	EL_{true}	$a^{EL,25}_{2.5\%}$	EL^{25}	$b^{EL,25}_{97.5\%}$	$a^{EL,1000}_{2.5\%}$	EL^{1000}	$b^{EL,1000}_{97.5\%}$
CCC	14	29	0.4872	0.2714	0.4196	0.5678	0.4626	0.4870	0.5115
B	3	6	0.5155	0.2577	0.4136	0.5696	0.4798	0.5053	0.5309
BB	0	3	0.3617	0.1991	0.2941	0.3890	0.3409	0.3596	0.3783
BBB	0	3	0.0899	0.0458	0.0839	0.1221	0.0844	0.0905	0.0966
A	0	3	0.0133	0.0034	0.0107	0.0180	0.0112	0.0128	0.0144

TABLE 17. Simulation study $n = 25$ and $n = 1000$, $Y \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$: expected loss of CDO tranches mapped to moneyness, $\rho_j \neq \rho_k$

shows that the estimation technique would produce the same ρ if they were actually equal! Looking at Table 6, we could conclude that the correlation for CCC ratings is higher than for other ratings. This might be explained by the fact that a bad state of the economy influences bad ratings more than others. However, CCC rated company usually possess a high idiosyncratic risk, i.e. ρ should be lower. This can be explained by the following reason: the number of CCC companies in our data set is quite low in comparison to the other ratings, correlation might be biased by $N_{CCC,t}$. Further, high correlation is usually possessed by large companies. As A companies are more likely to be large, the correlation in A companies should be higher. A correlation study grouped by the rating seems to be questionable, not least as correlation hardly depends on the rating class rather on the specific industry or country. One should further notice that the empirical confidence intervals of p are wider within the simulation study with different input ρ. Our concept of estimating default probability and asset correlation in a row is therefore very important.

Unfortunately, we next recognize that confidence intervals for ρ are not very accurate for better rating classes, see e.g. Table 14. This is because $s^2_{\tilde{L}}$ fluctuates less for better ratings and so $\hat{\rho}_{BO}$ does less, although ρ is approximately on the same level for every rating class. To give an example for B, BB, BBB, A ratings $\hat{\rho}_{BO}$ is around the true value $\rho = 0.08$ but $s^2_{\tilde{L}}$ becomes a lot smaller with a better rating, see Table 14. Only a small change of $s^2_{\tilde{L}}$ causes a big difference in the estimator and the confidence bounds of ρ, see e.g. also Remark 3.2. We conclude that $s^2_{\tilde{L}}$ is not accurate enough to replace $V(g(Y))$. The inaccuracy of $V(g(Y))$ caused by the replacement with $s^2_{\tilde{L}}$ gets worse the better the rating, i.e. as p gets closer to 0. It is then carried forward to an inaccuracy in ρ and its confidence intervals. In theory this effect is balanced out for very large n. But even for $n = 1000$, this effect still appears. The reason is that the true values p and $V(g(Y))$ are close to the boundary 0, the approximation by a normal distribution of $m_{\tilde{L}}/s^2_{\tilde{L}}$ is still not very precise even for reasonable large n. Hence the normal approximation of $\hat{\rho}_{BO}$ is bad as well. Consequently, our confidence intervals, especially for better rating classes, do not contain the actually value. Fortunately, we see that the true expected loss always lies in our confidence intervals. However, the point estimate approaches a high quality for a high sample only, see Table 17. As already mentioned we should again think about estimating ρ only within rating classes. For some extensions see Bluhm and Overbeck (2003) and Niethammer (2007).

5. Concluding Remarks

In this paper we introduce a general concept of deriving estimated error bounds for general risk figures in a credit risk model with a one-period default probability p and a one-period asset correlation ρ. This procedure is independent of the special procedure of deriving parameter estimates. If distributional properties of the estimators are known, they can be carried forward to the distribution of smooth functions of the parameters. The delta method essentially does this job, see e.g. (van der Vaart, 1998). We can thus derive the distribution of risk figures in general, see Section 2.2.

Afterwards, we suggest several estimators for the Gaussian one factor model and prove that estimators are asymptotically normal as long as they do not lie on the boundary. Local asymptotic normality of the Gaussian factor model is proved. This enables us to define asymptotic efficiency of the parameter estimates. Asymptotic efficiency is then examined for all estimators. Maximum likelihood estimators are efficient. Although moment estimation is not efficient, finite sample properties seem to be a lot better. This is supported by a simulation study. We therefore look at moment estimation in detail. We derive confidence intervals for parameters and risk figures. An empirical study presents parameter estimates and actual risk figures with its confidence intervals on a one and five year horizon.

We go on with some related literature and possible extensions. Our estimators possess the properties of so called extremum estimators (ML-estimator, (general) method of moments etc.). We could therefore try to apply quite general results already proven for this class of estimators by proving general conditions, see Newey and McFadden (1994). However, for our special model our main asymptotic results also come out by a simple application of usual limit theory as the law of large numbers, the central limit theorem, and local minimax theorems. A good references are Andrews (1992, 1999); Serfling (1980); Ibragimov and Has'minskii (1981); van der Vaart (1998). So not to break a butterfly on a wheel, we sticked to a derivation by hand.

We have further seen that the Gaussian approximation of the distribution of the moment estimator $\hat{\rho}_{BO}$ is not good for better quality ratings, i.e. for buckets with a small default probability. This is because $\hat{\rho}_{BO}$ is driven by $m_{\tilde{L}}$ and $s^2_{\tilde{L}}$ which are very small and close to the boundary zero. So for both the normal approximation is extremely bad and holds for a very large n only. Hence, it cannot be perfect for $\hat{\rho}_{BO}$. This problem is dealt with in general in (Andrews, 1999). He explains what is to do when the boundary is hit. The distribution might not be normal any longer. Many other interesting examples are described. Finally for future research if it is possible to establish all necessary assumptions given in (Andrews, 1999), we could write down the distribution at the boundary of our moment estimator in a sense of an extremum estimator even for a non-i.i.d. common factor. Asymptotic confidence intervals could be derived. Moreover, confidence of the ML-estimator in an infinite granular (ML, $N_t \to \infty$) can be derived exactly as the distribution of $\tilde{L}_{t,k}$ is analytically given.

For the sake of completeness, another methodology to quantify confidence intervals is bootstrapping. However, confidence intervals can be also inconsistent, see (Andrews, 1999; Beran, 1997). Subsampling presents a quite developed alternative and preceded the bootstrap. For both methods, we refer to Beran (1997); Bickel et al. (1997); Politis (1999). Finally, we still do not know what is to do if no or almost no defaults appeared, e.g. facing a AAA bucket of firms. A possible solution is given by Pluto and Tasche (2005). An estimation principle called *most prudent estimation* is proposed.

References

Andrews, D. W. K. (1984), Non-strong mixing autoregressive processes, Journal of Applied Probability, 21, 930-934.

Andrews, D. W. K. (1992), Generic Uniform Convergence, Econometric Theory, 8, 241-257.

Andrews, D. W. K. (1999), Estimation when a Parameter is on a Boundary - Generic Uniform Convergence, Econometrica, 67, 6, 1341-1383.

Beran, R. (1997), Diagnosing Bootstrap Success, Annals of the Institute of Statistical Mathematics, 49, 1-24.

Bickel, P.J., Götze, F., and van Zwet, W. R. (1997), Resampling Fewer than n Observations: Gains, Losses, Remedies for Losses, Statistica Sinica, 7, 1-31.

Bluhm, C. and Overbeck, L. (2003), Estimating Systematic Risk in Uniform Credit Portfolios, Credit Risk. Hrsg.: G.Bol et al. Contributions to Economics Physica-Verlag, Heidelberg.

Black, F. and Scholes, M. (1973), The Pricing of Options and Corporate Liabilities; Journal of Political Economy 81, 637-654.

Bluhm, C., Overbeck, L., and Wagner, C.K.J. (2002), An Introduction to Credit Risk Modeling, Financial Mathematics Series, Chapman & Hall/CRC, London.

Brockwell, P. J. and Davis, R. A. (1987), Times Series: Theory and Methods, Springer Series in Statistics, Springer-Verlag, New York.

Le Cam and Yang (1990), Asymptotics in Statistics - Some Basic Concepts, Springer Series in Statistics, Springer-Verlag, New York.

Davidson, J. (1994), Stochastic Limit Theory, an Introduction for Econometricians, Oxford University Press Journal of Political Economy 81, 637-654.

Gössl, C. (2005), Predictions Based on Certian Uncertainties - a Bayesian Credit Portfolio Approach, HypoVereinsbank AG, London, preprint.

Ibragimov, I. A. (1962), Some limit theorems for stationary processes, Theory of Probability and its Applications, 7 349-382.

Ibragimov, I. A. and Has'minskii, R. Z. (1981), Statistical Estimation - Asymptotic Theory, Applied Mathematics, Springer-Verlag, New York.

Kalemanova, A., Schmid, B., and Werner, R. (2005), The Normal Inverse Gaussian Distribution for Synthetic CDO pricing, www.default-risk.com.

Lando, D. (2004), Credit Risk Modeling: Theory and Applications, Princeton Series in Finance, Princeton University Press.

Merton, R. C. (1974), On the Pricing of Corporate Debt: the Risk Structure of Interest Rates, Journal of Finance, 29, 449-470.

Newey, W. K. and McFadden, D. (1994), Large Sample Estimation and Hypothesis Testing, Elsevier Science, Handbook of Econometrics, Volume IV, Chapter 36, 2111-2245.

Niethammer, C. R. (2007), Are Default Correlations time-dependent? A Bayesian Approach, preprint.

Pluto, K. and Tasche, D. (2005), Estimating Probabilities of Default for Low Default Portfolios, http://www.defaultrisk.com/pp_score_45.htm.

Politis, D. N., Romano, J. P., and Wolf M. (1999), Subsampling, Springer Series in Statistics, Springer-Verlag Berlin.

Serfling, R. J. (1980), Approximation Theorems of Mathematical Statistics, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons.

S&P default study, Annual 2005 Global Corporate Default Study and Rating Transitions, Standard& Poor's, Global Fixed Income Research, January 2006.

van der Vaart, A.W. (1998), Asymptotic Statistics, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

Vasicek, O. A. (1991a), The Loan Loss Distribution, Technical Report, KMV Corporation.

Vasicek, O. A. (1991b), Limiting Loan Loss Probability Distribution, Technical Report, KMV Corporation.

Vasicek, O. A. (1996/98), A Series Expansion for the Bivariate Normal Integral, Technical Report, KMV Corporation.

E-mail address: `christina.niethammer@uni-konstanz.de` (University of Konstanz/UniCredit Markets&Investment Banking)

E-mail address: `ludger.overbeck@math.uni-giessen.de` (University of Giessen/Commerzbank)