TECHNISCHE UNIVERSITÄT
DRESDEN
Fakultät Wirtschaftswissenschaften

Dresdner Beiträge zu
Quantitativen Verfahren

Nr. 45/05

Predicting the Credit Cycle with an Autoregressive Model

von

Steffi Höse und Konstantin Vogl

Herausgeber:
Die Professoren der
Fachgruppe Quantitative Verfahren

ISSN 0945-4802

Predicting the Credit Cycle with an Autoregressive Model

Steffi Höse and Konstantin Vogl

August 2, 2005

Abstract

Credit default events show cross sectional as well as serial correlation. While the latter is often neglected by current credit risk models, this work incorporates both types of dependence. A Bernoulli mixture model is considered, where in each rating grade the probit of the stochastic Bernoulli parameter follows an autoregressive stationary process with exogenous variables. The model parameters are estimated for a large retail portfolio. Exemplarily, prediction intervals of the default probabilities of the best and worst non-default rating grade are given and predicted credit portfolio loss distributions are plotted in contrast to the unconditional loss distributions.

1 Motivation

One of the main problems in credit risk modeling is that the observations available for the estimation of model parameters are dependent over time. However, time series in the credit risk industry are often so short that calibrations are accomplished under the assumption of independence. Whereas this negligence could be accepted for a model fit, forecasts of default probabilities need to take the credit cycle into account. A parsimonious parameterization is required due to the scarce data situation. This paper examines what can be achieved using stationary autoregressive processes with an additional exogenous variable in order to incorporate the credit cycle.

2 Notations and assumptions

In the following survey a homogeneous portfolio with n_t obligors belonging to the same rating grade is considered over time periods $t = 1, \ldots, T$. Let $A_{t,i}$ denote the default-indicator variable of the i-th obligor in time period t, which can take a value of one in the case of default or zero otherwise. The defaults are modeled within a Bernoulli mixture model[1], where at the beginning of time period t, the stochastic default probability $\tilde{\pi}_t$ takes a number $\pi_t \in \,]0, 1[\,$.

The variables $A_{t,i}$ for $i = 1, 2, \ldots, n_t$ are assumed to be conditionally independent for given realizations of the stochastic default probabilities $\tilde{\pi}_t, \tilde{\pi}_{t-1}, \ldots$ and for given realizations of macroeconomic impact variables $V_{t-1}, V_{t-2}, \ldots$. Thus, their conditional distribution is a Bernoulli distribution,

$$A_{t,i} \Big| \tilde{\pi}_t = \pi_t, V_{t-1} = v_{t-1}, \tilde{\pi}_{t-1} = \pi_{t-1}, \ldots \quad \sim \quad \mathrm{Ber}(\pi_t), \tag{1}$$

[1]As an example see Joe (1997, p. 211) and Frey and McNeil (2003, p. 67).

with parameter π_t. Given these realizations, the default of each of the obligors during time period t occurs independently with probability π_t. Therefore,

$$P[A_{t,i} = 1 | \tilde{\pi}_t = \pi_t] = \pi_t,$$

is the conditional default probability during period t, and

$$P[A_{t,i} = 1] = E[\tilde{\pi}_t] \tag{2}$$

is the unconditional default probability.

The transformed[2] variables $\Phi^{-1}(\tilde{\pi}_t)$, in the sequel addressed as *probits*, are assumed to follow a strictly stationary first-order autoregressive process,

$$\Phi^{-1}(\tilde{\pi}_t) = \alpha + \beta\, \Phi^{-1}(\tilde{\pi}_{t-1}) + V_{t-1} + U_t, \tag{3}$$

where $\alpha \in \mathbb{R}$ and $-1 < \beta < 1$. The random variable V_{t-1} plays the role of an observable macroeconomic impact, whereas U_t includes the remaining irregular component. All the random variables V_t and U_t are assumed to be mutually independent and Gaussian distributed,

$$U_t \overset{i.i.d.}{\sim} \mathrm{N}(0, \sigma_U^2), \qquad V_t \overset{i.i.d.}{\sim} \mathrm{N}(\mu_V, \sigma_V^2). \tag{4}$$

Equation (3) represents an ARX-model[3] for the probits $\Phi^{-1}(\tilde{\pi}_t)$, where AR refers to the autoregressive part and X to the exogenous input variables V_t. Aside from these exogenous macroeconomic impact variables, the model presented here is equivalent to a Basel II single-factor model[4] with a stationary first-order autoregressive process for the systematic factor, as shown in section 8.5.

3 Basic moments

In this framework the model is parameterized by $\alpha, \beta, \sigma_U^2, \mu_V$ and σ_V^2. The previously made assumptions lead to Gaussian distributed probits

$$\Phi^{-1}(\tilde{\pi}_t) \sim \mathrm{N}\left(\frac{\alpha + \mu_V}{1-\beta}, \frac{\sigma_U^2 + \sigma_V^2}{1-\beta^2}\right), \tag{5}$$

which are stationary and autocorrelated with k-th order autocorrelation β^k. As a consequence, the stochastic default probabilities $\tilde{\pi}_t$ are also dependent over time and are stationary Vasicek[5] distributed,

$$\tilde{\pi}_t \sim \mathrm{Vas}(\pi, \varrho), \tag{6}$$

with parameters $\pi \in (0,1)$ and $\varrho \in (0,1)$, which can also be written as functions of the model parameters. The expectation is determined by

$$\pi := E[\tilde{\pi}_t] \tag{7}$$

$$= \Phi\left(\frac{(\alpha + \mu_V)\sqrt{1-\beta^2}}{(1-\beta)\sqrt{1-\beta^2+\sigma_U^2+\sigma_V^2}}\right), \tag{8}$$

[2]The notation Φ^{-1} is used for the inverse of the cumulative distribution function of the standardized Gaussian distribution.

[3]See e. g. Ljung (1999).

[4]See Basel Committee on Banking Supervision (2004).

[5]The definition of the Vasicek distribution (or Probit-normal distribution (Frey and McNeil, 2003)) is given in section 8.

which plays the role of the unconditional default probability in the rating grade, see (2). The variance is given by

$$V[\tilde{\pi}_t] = \Phi_2\left(\Phi^{-1}(\pi), \Phi^{-1}(\pi); \varrho\right) - \pi^2, \tag{9}$$

where $\Phi_2(\cdot, \cdot; \varrho)$ denotes the cumulative distribution function of the standardized bivariate Gaussian distribution with the correlation parameter

$$\varrho := \frac{\sigma_U^2 + \sigma_V^2}{1 - \beta^2 + \sigma_U^2 + \sigma_V^2}\,. \tag{10}$$

The proofs of all equations are given in section 8.

4 Distribution of default indicators

The Bernoulli mixture model given in (1) implies that the default indicators are Bernoulli distributed random variables,

$$A_{t,i} \sim \mathrm{Ber}(\pi), \tag{11}$$

with default probability π from equation (8). Using the conditional independence of $A_{t,i}$ and $A_{t,j}$ for $i \neq j$ the product moment can be written as

$$E[A_{t,i}A_{t,j}] = E[\tilde{\pi}_t^2] = V[\tilde{\pi}_t] + (E[\tilde{\pi}_t])^2$$

and therefore the covariance yields to

$$Cov[A_{t,i}, A_{t,j}] = V[\tilde{\pi}_t].$$

Thus, the default indicators within the rating grade are equicorrelated with a non-negative correlation,

$$\begin{aligned} Corr[A_{t,i}, A_{t,j}] &= \frac{V[\tilde{\pi}_t]}{\pi(1-\pi)} \\ &= \frac{\Phi_2\left(\Phi^{-1}(\pi), \Phi^{-1}(\pi); \varrho\right) - \pi^2}{\pi(1-\pi)}\,, \end{aligned} \tag{12}$$

which shows the relation between the correlation of the default indicators and ϱ.

With lagged variables as the condition, the default indicators are also Bernoulli distributed,

$$A_{t,i}|\tilde{\pi}_{t-1} = \pi_{t-1}, V_{t-1} = v_{t-1} \quad \sim \quad \mathrm{Ber}(p_t), \tag{13}$$

where the Bernoulli parameter

$$p_t := \Phi\left(\frac{\alpha + \beta\,\Phi^{-1}(\pi_{t-1}) + v_{t-1}}{\sqrt{1+\sigma_U^2}}\right) \tag{14}$$

additionally depends on the realizations of the default probability and the macroeconomic impact in period $t-1$. The conditional covariance of the default indicators is equal to the conditional variance of the stochastic default probability, which is determined by

$$Cov[A_{t,i}, A_{t,j}|\tilde{\pi}_{t-1}, V_{t-1}] = V[\tilde{\pi}_t|\tilde{\pi}_{t-1}, V_{t-1}], \qquad \forall i \neq j, \tag{15}$$

$$V[\tilde{\pi}_t|\tilde{\pi}_{t-1} = \pi_{t-1}, V_{t-1} = v_{t-1}] = \Phi_2\left(\Phi^{-1}(p_t), \Phi^{-1}(p_t); \tfrac{\sigma_U^2}{1+\sigma_U^2}\right) - p_t^2, \tag{16}$$

with p_t from (14).

5 Prediction of default probabilities

The autoregressive model of dependence over time can be used to predict the stochastic default probability on the basis of the past observations. The model discussed in the sections 2 to 4 is applied to a hypothetical portfolio of n_t retail clients in a single rating grade. In order to estimate the probit-AR(1)-process, an authentic data set from the SCHUFA[6] is used, which contains the default history of 16 quarterly observations from I/2000 to IV/2003 of about 800 000 German retail clients. The macroeconomic impact variable is chosen to be

$$V_t = \gamma_1 X_t^{(1)} + \gamma_2 X_t^{(2)}, \tag{17}$$

where the exogenous variable $X_t^{(1)}$ denotes the change of the logarithm of the disposable income of German households, and the exogenous variable $X_t^{(2)}$ denotes the change of the logarithm of the German unemployment rate.[7]

5.1 Probit-AR(1)-process with macroeconomic impacts

Under the assumption of the probit-AR(1)-process of equation (3) and (4) the conditional distribution of the probits is a Gaussian distribution,

$$\Phi^{-1}(\tilde{\pi}_t)|\tilde{\pi}_{t-1} = \pi_{t-1}, V_{t-1} = v_{t-1} \sim \mathrm{N}(\alpha + \beta\Phi^{-1}(\pi_{t-1}) + v_{t-1}, \sigma_U^2). \tag{18}$$

Therefore, the $1-\alpha^*$ prediction interval for $\tilde{\pi}_t$ is given by $[z_l, z_u]$, where

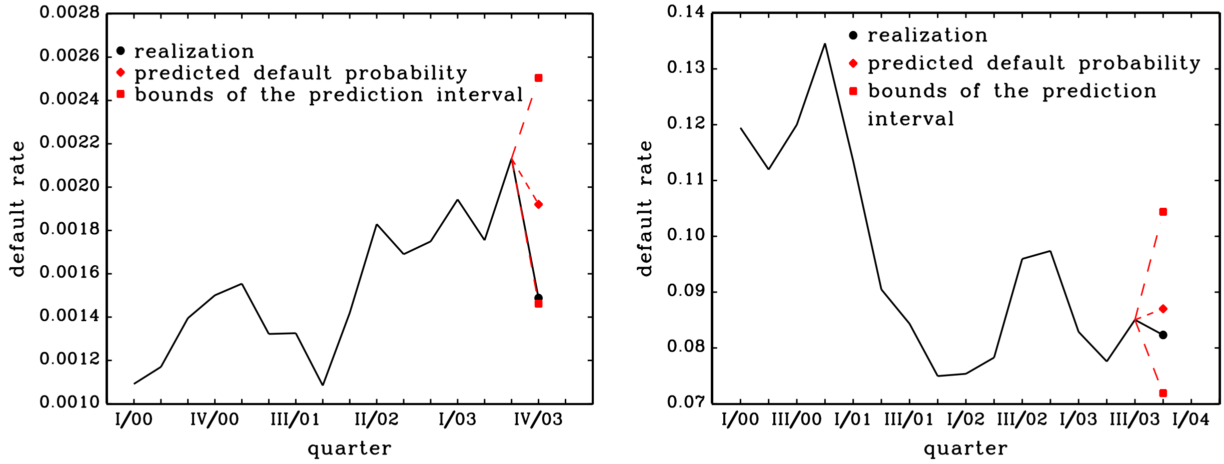
$$\begin{aligned} z_l &:= \Phi\left(\alpha + \beta\,\Phi^{-1}(\pi_{t-1}) + v_{t-1} - \sigma_U\,\Phi^{-1}(1-\tfrac{\alpha^*}{2})\right), \\ z_u &:= \Phi\left(\alpha + \beta\,\Phi^{-1}(\pi_{t-1}) + v_{t-1} + \sigma_U\,\Phi^{-1}(1-\tfrac{\alpha^*}{2})\right). \end{aligned}$$

If the task was to predict $\Phi^{-1}(\tilde{\pi}_t)$ given $\tilde{\pi}_{t-1} = \pi_{t-1}, V_{t-1} = v_{t-1}$, the prediction would clearly be $\alpha + \beta\Phi^{-1}(\pi_{t-1}) + v_{t-1}$, which is the mean and the median of the predicted distribution in (18). But if one is interested in predicting $\tilde{\pi}_t$ given $\tilde{\pi}_{t-1} = \pi_{t-1}, V_{t-1} = v_{t-1}$, the mean and the median of the predicted distribution differ. Whereas the mean is given by (14), the median is $\Phi\left(\alpha + \beta\,\Phi^{-1}(\pi_{t-1}) + v_{t-1}\right)$ which is used in Figure 1 in order to predict the default probability of rating grade A (highest creditworthiness) and rating grade M (worst non-default grade). The observed default rates from I/2000 to III/2003 are displayed as solid lines and the predicted default probabilities of the last quarter IV/2003 are shown as diamonds. The squares represent the bounds of the 95% prediction intervals around the point estimator. Thereby, the model parameters are estimated from 15 observations (I/2000 to IV/2003). In order to ensure stationarity, the parameter β is estimated by the autocorrelation of time lag one. In a further step the parameters α, γ_1, γ_2 and σ_U^2 are estimated by the ordinary least squares method. The observed default rate in quarter III/2003 estimates the realization π_{t-1}. With observed values of the exogenous variables $X_{t-1}^{(1)}$ and $X_{t-1}^{(2)}$ in III/2003 and with the least-squares-estimates of γ_1 and γ_2 an estimate of v_{t-1} in III/2003 is obtained according to (17). The parameters μ_V and σ_V^2 are estimated as sample mean and sample variance of the macroeconomic impact variables V_t given in (17), when the estimates of γ_1 and γ_2 are known.

The model parameters were estimated as follows.

[6] The SCHUFA AG is one of the major suppliers of consumer credit scores in Germany, comparable to EXPERIAN, EQUIFAX or TRANS UNION in the USA.

[7] $X_t^{(1)} = \ln Y_t^{(1)} - \ln Y_{t-1}^{(1)}$ and $X_t^{(2)} = \ln Y_t^{(2)} - \ln Y_{t-1}^{(2)}$, where for $Y^{(1)}$ the time series BDJA9405B and for $Y^{(2)}$ the time series BDOUN013R from DATASTREAM are used.



(a) Rating grade A (b) Rating grade M

Figure 1: Predicted default probabilities and default rates in a probit-AR(1)-process

parameter	estimate for rating grade A	rating grade M
α	-1.3024	-0.3372
β	0.5532	0.7415
σ_U^2	0.001861	0.002734
σ_V^2	0.0003161	0.001825
μ_V	-0.01195	-0.01460

In contrast to rating grade A the quality of prediction in rating grade M is fairly good despite the fact that the model is only capable of reacting to recent changes with a delay of one period.

5.2 Probit-AR(2)-process with macroeconomic impacts

In this section an enhanced model is considered, which incorporates more than one time lag. For that reason assumption (3) is replaced by

$$\Phi^{-1}(\tilde{\pi}_t) = \alpha + \beta_1\,\Phi^{-1}(\tilde{\pi}_{t-1}) + \beta_2\,\Phi^{-1}(\tilde{\pi}_{t-2}) + V_{t-1} + U_t,$$

where $\alpha \in \mathbb{R}$ and $(\beta_1, \beta_2) \in \mathbb{R}^2$. Again the random variables V_t and U_t are assumed to be mutually independent and Gaussian distributed, see (4). This means a second-order autoregressive process is assumed for the probits of the stochastic default probabilities. To ensure that this process is stationary over time the following conditions have to be fulfilled

$$\begin{aligned} \beta_1 + \beta_2 &< 1, \\ \beta_2 - \beta_1 &< 1, \\ \beta_2 &> -1. \end{aligned}$$

Under these modified assumptions the probits of the default probabilities are also Gaussian distributed,

$$\Phi^{-1}(\tilde{\pi}_t) \sim \mathrm{N}\left(\frac{\alpha + \mu_V}{1 - \beta_1 - \beta_2}, \frac{\sigma_U^2 + \sigma_V^2}{1 - \beta_1^2 - \beta_2^2 - 2\beta_1\beta_2\frac{\beta_1}{1-\beta_2}}\right),$$

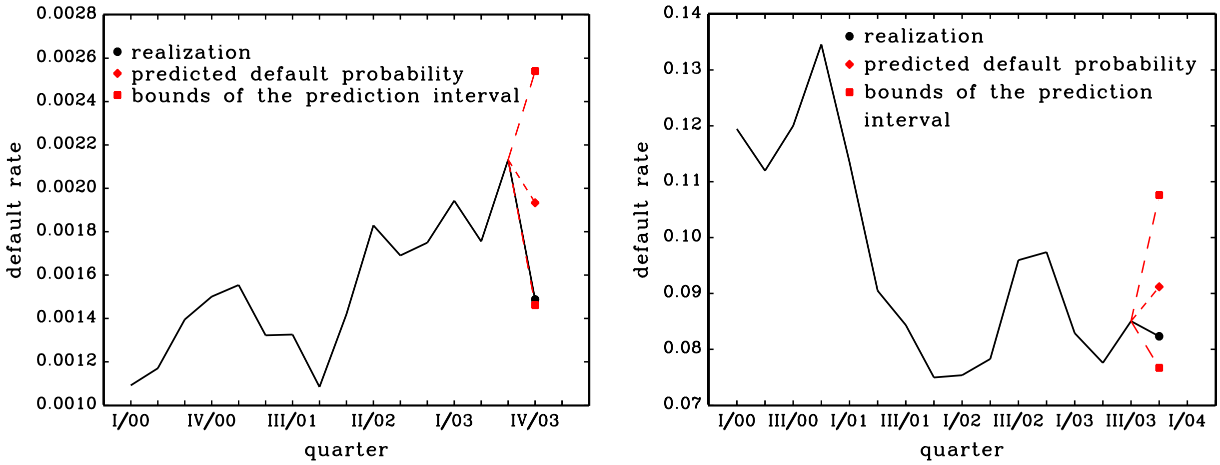
where mean and variance of the stationary distribution are functions of the five model parameters $\alpha, \beta_1, \beta_2, \sigma_U^2, \mu_V$ and σ_V^2.[8]

In this model the conditional distribution of the probits is also a Gaussian distribution,

$$\begin{aligned}\Phi^{-1}(\tilde{\pi}_t)|\tilde{\pi}_{t-1} = \pi_{t-1}, V_{t-1} = v_{t-1}, \pi_{t-2} = \pi_{t-2} \\ \sim \mathrm{N}(\alpha + \beta_1\Phi^{-1}(\pi_{t-1}) + \beta_2\Phi^{-1}(\pi_{t-2}) + v_{t-1}, \sigma_U^2),\end{aligned}$$

and the $1-\alpha^*$ prediction interval for $\tilde{\pi}_t$ is given by $[z_l, z_u]$, where

$$\begin{aligned} z_l &:= \Phi\left(\alpha + \beta_1\,\Phi^{-1}(\pi_{t-1}) + \beta_2\,\Phi^{-1}(\pi_{t-2}) + v_{t-1} - \sigma_U\,\Phi^{-1}(1-\tfrac{\alpha^*}{2})\right), \\ z_u &:= \Phi\left(\alpha + \beta_1\,\Phi^{-1}(\pi_{t-1}) + \beta_2\,\Phi^{-1}(\pi_{t-2}) + v_{t-1} + \sigma_U\,\Phi^{-1}(1-\tfrac{\alpha^*}{2})\right). \end{aligned}$$



(a) Rating grade A

(b) Rating grade M

Figure 2: Predicted default probabilities and default rates in a probit-AR(2)-process

The quality of predictions is demonstrated in Figure 2. Again, in order to ensure stationarity, the parameters β_1 and β_2 are estimated via the autocorrelation according to the Yule-Walker equations

$$\begin{aligned}\beta_1 &= \frac{\varrho_1(1-\varrho_2)}{1-\varrho_1^2}, \\ \beta_2 &= \frac{\varrho_2-\varrho_1^2}{1-\varrho_1^2},\end{aligned}$$

where ϱ_i is the autocorrelation of time lag i. Subsequently, the parameters α, γ_1, γ_2 and σ_U^2 are estimated by the ordinary least squares method, where now only 14 observations from II/2000

[8]Therefore, the stochastic default probabilities remain Vasicek distributed, $\tilde{\pi}_t \sim \mathrm{Vas}(\pi, \varrho)$, but the parameters $\pi \in]0,1[$ and $\varrho \in]0,1[$ differ from the parameters of the probit-AR(1)-model. So equations (8) and (10) have to be replaced by

$$\pi := \Phi\left(\frac{(\alpha+\mu_V)\sqrt{1-\beta_1^2-\beta_2^2-2\beta_1\beta_2\frac{\beta_1}{1-\beta_2}}}{(1-\beta_1-\beta_2)\sqrt{1-\beta_1^2-\beta_2^2-2\beta_1\beta_2\frac{\beta_1}{1-\beta_2}+\sigma_U^2+\sigma_V^2}}\right),$$

and

$$\varrho := \frac{\sigma_U^2+\sigma_V^2}{1-\beta_1^2-\beta_2^2-2\beta_1\beta_2\frac{\beta_1}{1-\beta_2}+\sigma_U^2+\sigma_V^2}.$$

to III/2003 can be used. The realizations π_{t-1} and π_{t-2} are estimated by the default rates observed in III/2003 and II/2003.

Now the model parameters were estimated as follows.

	estimate for	
parameter	rating grade A	rating grade M
α	-1.3389	-0.4650
β_1	0.5692	1.0130
β_2	-0.02882	-0.3662
σ_U^2	0.001969	0.002308
σ_V^2	0.0003137	0.001802
μ_V	-0.01137	-0.005201
ϱ_1	0.5532	0.7415

However, the inflation to a probit-AR(2)-model is not able to show striking improvements compared to the probit-AR(1)-model. There is also no evidence to reject the hypothesis that the second-order partial autocorrelation coefficient, which is equal to β_2 in the AR(2)-model, equals zero.[9] But there may be applications where the AR(2)-model shows significant advantages, especially when longer time series are available. In contrast, the first-oder partial autocorrelation coefficient, which equals ϱ_1, differs more than two standard deviations from zero in both rating grades.

6 Credit portfolio loss distribution

The portfolio loss incurred in period t is defined as[10]

$$L_t := \sum_{i=1}^{n_t} w_{t,i} A_{t,i},$$

where $w_{t,i} \geq 0$ is the product of the exposure at default and the loss given default caused by the i-th obligor during period t.

6.1 Moments of the loss distribution

The following remarks are made within the framework of a probit-AR(2)-model.[11] The unconditional expected loss EL and the stochastic conditional expected loss cEL_t in period t are defined by

$$\mathsf{EL} := E[L_t] = \pi \sum_{i=1}^{n_t} w_{t,i} \tag{19}$$

and

$$\mathsf{cEL}_t := E[L_t|\tilde{\pi}_{t-1}, V_{t-1}, \tilde{\pi}_{t-2}] = P_t \sum_{i=1}^{n_t} w_{t,i}, \tag{20}$$

where the random variable

$$P_t := \Phi\left(\frac{\alpha + \beta_1\,\Phi^{-1}(\tilde{\pi}_{t-1}) + \beta_2\,\Phi^{-1}(\tilde{\pi}_{t-2}) + V_{t-1}}{\sqrt{1+\sigma_U^2}}\right) \tag{21}$$

[9] Considering a time series of 16 observations, the standard deviation of the partial autocorrelation coefficients is approximately 0.25, cf. Anderson (1976, p. 9).

[10] See Basel Committee on Banking Supervision (2004).

[11] The probit-AR(1)-model can be considered for $\beta_2 = 0$.

is defined analogously to the realized conditional default probability of the probit-AR(1)-model, which is given in (14). In order to measure the uncertainty of the portfolio loss, the unexpected loss[12] UL can be defined by the square root of

$$\mathsf{UL}^2 := V[L_t] = \pi(1-\pi)\sum_{i=1}^{n_t} w_{t,i}^2 + V[\tilde{\pi}_t]\sum_{\substack{i,j=1\\ i\neq j}}^{n_t} w_{t,i}w_{t,j} \tag{22}$$

in the unconditional case and by the square root of

$$\begin{aligned}\mathsf{cUL}_t^2 &:= V[L_t|\tilde{\pi}_{t-1}, V_{t-1}, \tilde{\pi}_{t-2}] \\ &= P_t(1-P_t)\sum_{i=1}^{n_t} w_{t,i}^2 + V[\tilde{\pi}_t|\tilde{\pi}_{t-1}, V_{t-1}, \tilde{\pi}_{t-2}]\sum_{\substack{i,j=1\\ i\neq j}}^{n_t} w_{t,i}w_{t,j}\end{aligned} \tag{23}$$

when the conditional distribution is considered. If an AR(2)-process is considered, the unconditional variance $V[\tilde{\pi}_t]$ is defined in (9) with π and ϱ from footnote 8 and the conditional variance

$$V[\tilde{\pi}_t|\tilde{\pi}_{t-1}, V_{t-1}, \tilde{\pi}_{t-2}] = \Phi_2\left(\Phi^{-1}(P_t), \Phi^{-1}(P_t); \tfrac{\sigma_U^2}{1+\sigma_U^2}\right) - P_t^2$$

is defined analogously to that given in equation (16) using P_t from equation (21).

Whereas the random variable cEL_t scatters around the unconditional expected loss EL,

$$E[\mathsf{cEL}_t] = \mathsf{EL},$$

the case is somewhat different if unexpected losses are considered. Using the law of total variance

$$V[L_t] = E[V[L_t|\tilde{\pi}_{t-1}, V_{t-1}, \tilde{\pi}_{t-2}]] + V[E[L_t|\tilde{\pi}_{t-1}, V_{t-1}, \tilde{\pi}_{t-2}]]$$

and Jensen's inequality one can show that the expectation of the conditional unexpected loss is equal to or less than the unconditional unexpected loss,

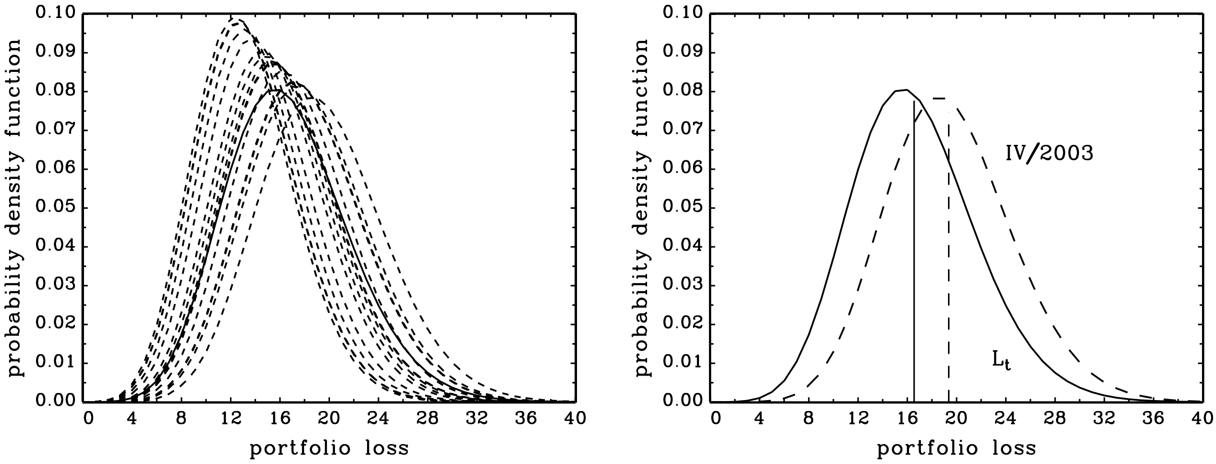
$$E[\mathsf{cUL}_t] \leq \sqrt{E[\mathsf{cUL}_t^2]} = \sqrt{E[V[L_t|\tilde{\pi}_{t-1}, V_{t-1}, \tilde{\pi}_{t-2}]]} \leq \sqrt{V[L_t]} = \mathsf{UL}. \tag{24}$$

This suggests that the unexpected loss could be reduced by conditioning. However, it is emphasized that this holds only in the mean. These results are also valid within the probit-AR(1)-model.

6.2 Prediction of the loss distribution

The conditional as well as the unconditional credit portfolio loss distribution is a mixture of Binomial distributions. Using the parameter estimates from the probit-AR-models, the loss distribution can be calculated by means of Monte Carlo methods. In Figures 3 and 4 the loss distributions of rating grade A and M with weights $w_{t,i} = 1$ for $n_t = 10\,000$ obligors are plotted within the probit-AR(1)-model. In 3(a) and 4(a) the unconditional and the conditional credit portfolio loss distributions are plotted as solid and dashed curves. In 3(b) and 4(b) the predicted distribution of the credit portfolio loss for IV/2003 is shown with long dashes, which is the conditional loss distributions when $\tilde{\pi}_{\text{III}/2003} = \pi_{\text{III}/2003}, V_{\text{III}/2003} = v_{\text{III}/2003}$ is given. The unconditional distribution is represented again by a solid curve. Additionally, the vertical lines mark the expected losses for each distribution. Analogously, Figures 5 and 6 show the loss distributions of rating grade A and M ($w_{t,i} = 1$ for $n_t = 10\,000$ obligors) when a probit-AR(2)-model is assumed. The predicted distribution of the credit portfolio loss for IV/2003 is the loss distribution given $\tilde{\pi}_{\text{III}/2003} = \pi_{\text{III}/2003}, V_{\text{III}/2003} = v_{\text{III}/2003}, \tilde{\pi}_{\text{II}/2003} = \pi_{\text{II}/2003}$.

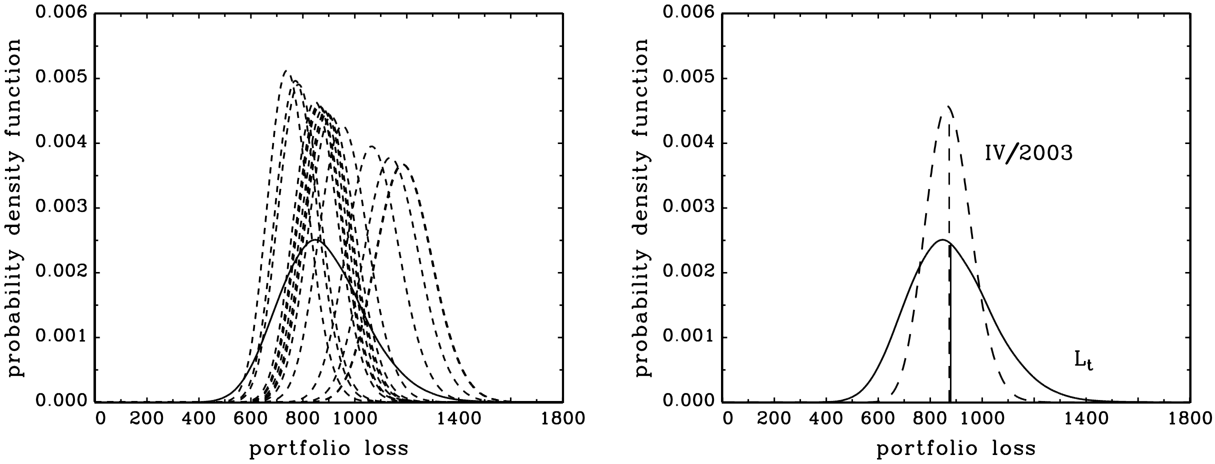
[12] The meaning of *unexpected loss* differs among authors. The interpretation here follows Bluhm et al. (2003, p. 28).



(a) Unconditional (solid) and conditional (dashed) portfolio loss distributions

(b) Unconditional (solid) and predicted loss distribution for IV/2003 (long dashed) with expected losses (vertical lines)

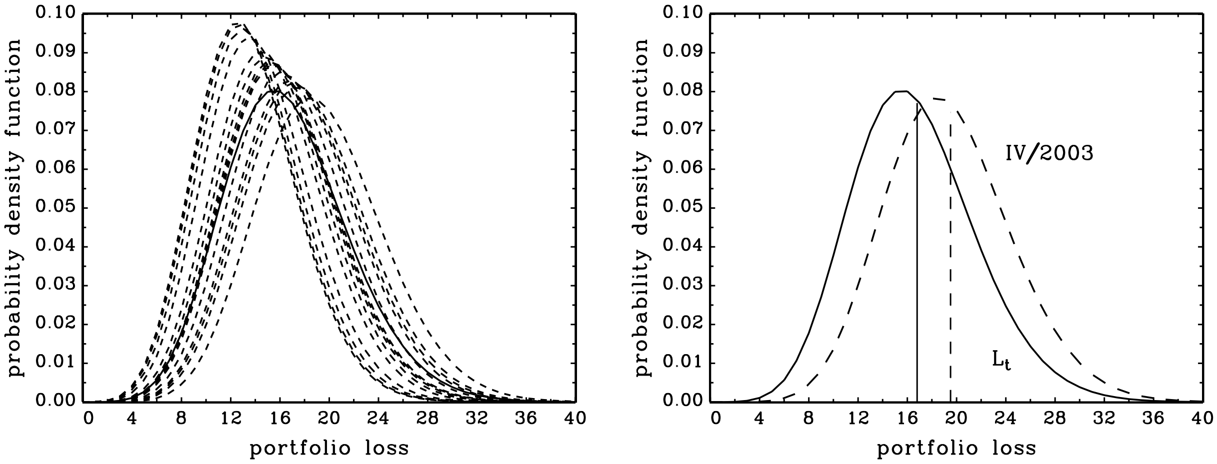
Figure 3: Credit portfolio loss distributions in a probit-AR(1)-model with $n_t = 10\,000$ rating grade A obligors with $w_{t,i} = 1$ for $i = 1, 2, \ldots, n_t$



(a) Unconditional (solid) and conditional (dashed) portfolio loss distributions

(b) Unconditional (solid) and predicted loss distribution for IV/2003 (long dashed) with expected losses (vertical lines)

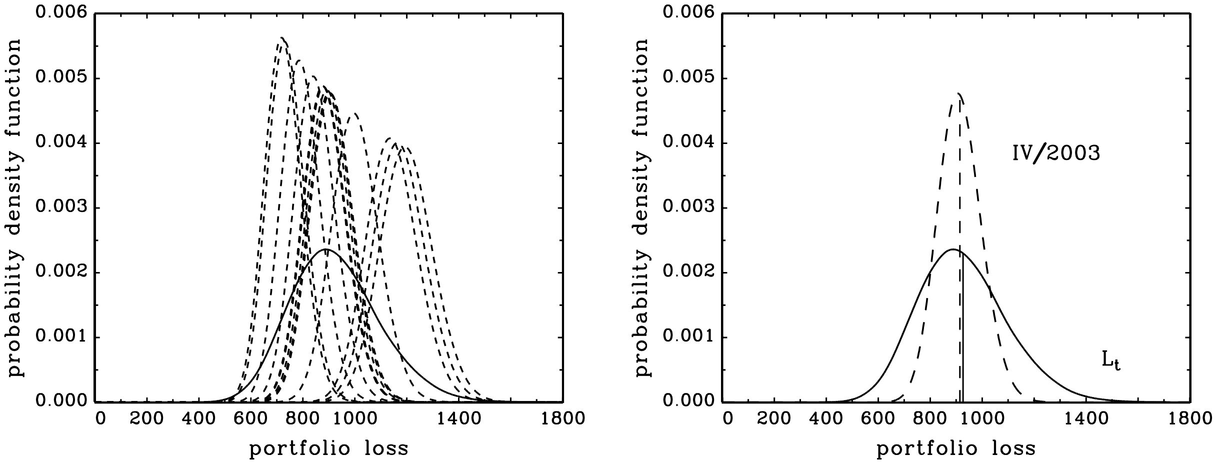
Figure 4: Credit portfolio loss distributions in a probit-AR(1)-model with $n_t = 10\,000$ rating grade M obligors with $w_{t,i} = 1$ for $i = 1, 2, \ldots, n_t$



(a) Unconditional (solid) and conditional (dashed) portfolio loss distributions

(b) Unconditional (solid) and predicted loss distribution for IV/2003 (long dashed) with expected losses (vertical lines)

Figure 5: Credit portfolio loss distributions in a probit-AR(2)-model with $n_t = 10\,000$ rating grade A obligors with $w_{t,i} = 1$ for $i = 1, 2, \ldots, n_t$



(a) Unconditional (solid) and conditional (dashed) portfolio loss distributions

(b) Unconditional (solid) and predicted loss distribution for IV/2003 (long dashed) with expected losses (vertical lines)

Figure 6: Credit portfolio loss distributions in a probit-AR(2)-model with $n_t = 10\,000$ rating grade M obligors with $w_{t,i} = 1$ for $i = 1, 2, \ldots, n_t$

7 Conclusions

If one obtaines a very extreme realization for the variables $\tilde{\pi}_{t-1}$ and V_{t-1} which condition the predicted loss distribution of period t, then the prediction tends to be an extreme one too. As a consequence, the default rate of rating grade A for IV/2003 nearly misses the 95% prediction intervals shown in Figures 1(a) and 2(a). In such situations with an extremely high probability of default one can see from Figures 3(b) and 5(b) that the conditional unexpected loss can be greater than the unconditional unexpected loss as emphasized at the end of section 6.1.

Comparing the Figures 1, 3 and 4 generated within the AR(1)-model to the Figures 2, 5 and 6 generated within the AR(2)-model, the differences seem to be minor ones. This can be interpreted as endorsement of the AR(1)-model, because even the more general approach yields almost the same predicted distributions. The question whether the AR(2)-model approach turns out to be superior can only be answered when longer time series are available.

If model parameters are estimated by the ordinary least squares method, it is possible that the estimates of β_1 and β_2 do not fulfill the stationarity conditions. Therefore, the described two step estimation method is used. In Höse and Vogl (2005) however, where only the AR(1)-model is investigated, the parameter estimates drawn from a one step OLS estimation never conflicted with the stationarity conditions.

8 Appendix

For the sake of simplicity the following definition is made.

Definition 1. *A random variable X with $P[0 < X < 1] = 1$ is called to be Vasicek distributed,*

$$X \sim \mathrm{Vas}(\pi, \varrho),$$

with parameters $\pi \in]0,1[$ and $\varrho \in]0,1[$, if

$$\Phi^{-1}(X) \sim \mathrm{N}\left(\frac{\Phi^{-1}(\pi)}{\sqrt{1-\varrho}}, \frac{\varrho}{1-\varrho}\right).$$

For $X \sim \mathrm{Vas}(\pi, \varrho)$ the properties

$$E[X] \;=\; \pi \tag{25}$$

and

$$V[X] \;=\; \Phi_2\left(\Phi^{-1}(\pi), \Phi^{-1}(\pi); \varrho\right) - \pi^2 \tag{26}$$

can be shown as follows.

8.1 Proof of the equations (25) and (26)

For $V, W, Z \overset{i.i.d.}{\sim} \mathrm{N}(0,1)$ and under the consideration of Definition 1, it holds true that

$$\Phi\left(\frac{\Phi^{-1}(\pi) - \sqrt{\varrho}Z}{\sqrt{1-\varrho}}\right) \sim \mathrm{Vas}(\pi, \varrho).$$

The expected value is then given by

$$\begin{aligned} E\left[\Phi\left(\frac{\Phi^{-1}(\pi) - \sqrt{\varrho}Z}{\sqrt{1-\varrho}}\right)\right] &= \\ &= \int_{-\infty}^{\infty} \Phi\left(\frac{\Phi^{-1}(\pi) - \sqrt{\varrho}z}{\sqrt{1-\varrho}}\right) d\Phi(z) \\ &= \int_{-\infty}^{\infty} P\left[\sqrt{\varrho}Z + \sqrt{1-\varrho}V \leq \Phi^{-1}(\pi) \middle| Z = z\right] d\Phi(z) \\ &= P\left[\sqrt{\varrho}Z + \sqrt{1-\varrho}V \leq \Phi^{-1}(\pi)\right] \;=\; \pi, \end{aligned}$$

which proves equation (25). Similarly it holds true that

$$\begin{aligned} E\left[\Phi\left(\frac{\Phi^{-1}(\pi) - \sqrt{\varrho}Z}{\sqrt{1-\varrho}}\right)^2\right] &= \\ &= \int_{-\infty}^{\infty} \Phi\left(\frac{\Phi^{-1}(\pi) - \sqrt{\varrho}z}{\sqrt{1-\varrho}}\right)^2 d\Phi(z) \\ &= \int_{-\infty}^{\infty} P\left[\sqrt{\varrho}Z + \sqrt{1-\varrho}V \leq \Phi^{-1}(\pi) \middle| Z = z\right] \\ &\qquad P\left[\sqrt{\varrho}Z + \sqrt{1-\varrho}W \leq \Phi^{-1}(\pi) \middle| Z = z\right] d\Phi(z) \\ &= \Phi_2\left(\Phi^{-1}(\pi), \Phi^{-1}(\pi); \varrho\right) \end{aligned}$$

and therefore

$$V\left[\Phi\left(\frac{\Phi^{-1}(\pi) - \sqrt{\varrho}Z}{\sqrt{1-\varrho}}\right)\right] = \Phi_2\left(\Phi^{-1}(\pi), \Phi^{-1}(\pi); \varrho\right) - \pi^2.$$

8.2 Proof of the equations (7) to (10)

From the assumption of the probit-AR(1)-process, which is defined by equations (3) and (4), it follows that the probits $\Phi^{-1}(\tilde{\pi}_t)$ are Gaussian distributed as can be seen in (5). Using Definition 1 one can show that $\tilde{\pi}_t$ has a specific Vasicek distribution as given in (6), where the parameters π and ϱ are the solution of the equations

$$\frac{\alpha + \mu_V}{1-\beta} = \frac{\Phi^{-1}(\pi)}{\sqrt{1-\varrho}} \qquad \text{and} \qquad \frac{\sigma_U^2 + \sigma_V^2}{1-\beta^2} = \frac{\varrho}{1-\varrho}$$

which is given in (8) and (10). From section 8.1 follow equations (7) and (9).

8.3 Proof of equations (2) and (12)

It follows from equation (1) that

$$A_{t,i}|\tilde{\pi}_t = \pi_t \overset{i.i.d.}{\sim} \text{Ber}(\pi_t)$$

for $i = 1, 2, \ldots, n_t$. With

$$\begin{aligned} E[A_{t,i}] &= P[A_{t,i} = 1] \\ &= \int_{[0,1]} P[A_{t,i} = 1|\tilde{\pi}_t = \pi_t]\, dF_{\tilde{\pi}_t}(\pi_t) \\ &= \int_{[0,1]} \pi_t\, dF_{\tilde{\pi}_t}(\pi_t) \\ &= E[\tilde{\pi}_t] \end{aligned}$$

equation (2) is proven. In this context $F_{\tilde{\pi}_t}$ is the cumulative distribution function of the stochastic default probability $\tilde{\pi}_t$.

Using the Bernoulli mixture model, particularly the independence of the random variables $A_{t,i}|\tilde{\pi}_t$ and $A_{t,j}|\tilde{\pi}_t$, and applying the law of iterated expectation the product moment for $i \neq j$ is given by

$$\begin{aligned} E[A_{t,i}A_{t,j}] &= E[E[A_{t,i}A_{t,j}|\tilde{\pi}_t]] \\ &= E[E[A_{t,i}|\tilde{\pi}_t]E[A_{t,j}|\tilde{\pi}_t]] \\ &= E[\tilde{\pi}_t^2] \\ &= V[\tilde{\pi}_t] + (E[\tilde{\pi}_t])^2. \end{aligned}$$

Now the covariance can be written as

$$\begin{aligned} Cov[A_{t,i}, A_{t,j}] &= E[A_{t,i}A_{t,j}] - E[A_{t,i}]E[A_{t,j}] \\ &= V[\tilde{\pi}_t] + (E[\tilde{\pi}_t])^2 - (E[\tilde{\pi}_t])^2 \\ &= V[\tilde{\pi}_t] \end{aligned}$$

and by using (11) the correlation of the default indicators for $i \neq j$ is given by

$$Corr[A_{t,i}, A_{t,j}] = \frac{Cov[A_{t,i}, A_{t,j}]}{\sqrt{V[A_{t,i}]V[A_{t,j}]}} = \frac{V[\tilde{\pi}_t]}{\pi(1-\pi)}.$$

8.4 Proof of equations (13) to (16)

It follows from equation (1) that the random variables $A_{t,i}$ given $\tilde{\pi}_{t-1} = \pi_{t-1}, V_{t-1} = v_{t-1}$ are Bernoulli distributed with parameter $E[A_{t,i}|\tilde{\pi}_{t-1} = \pi_{t-1}, V_{t-1} = v_{t-1}]$, which can be derived from

$$\begin{aligned} E[A_{t,i}|\tilde{\pi}_{t-1}, V_{t-1}] &= E[E[A_{t,i}|\tilde{\pi}_t, \tilde{\pi}_{t-1}, , V_{t-1}]|\tilde{\pi}_{t-1}, V_{t-1}] \\ &= E[\tilde{\pi}_t|\tilde{\pi}_{t-1}, V_{t-1}] \\ &= E\left[\Phi\big(\alpha + \beta\Phi^{-1}(\tilde{\pi}_{t-1}) + V_{t-1} + U_t\big)\big|\tilde{\pi}_{t-1}, V_{t-1}\right] \\ E[A_{t,i}|\tilde{\pi}_{t-1} = \pi_{t-1}, V_{t-1} = v_{t-1}] &= E\left[\Phi\big(\alpha + \beta\Phi^{-1}(\pi_{t-1}) + v_{t-1} + U_t\big)\right]. \end{aligned} \tag{27}$$

Due to $U_t \sim \mathrm{N}(0, \sigma_U^2)$ the variable $\Phi\big(\alpha + \beta\Phi^{-1}(\pi_{t-1}) + v_{t-1} + U_t\big)$ follows a specific Vasicek distribution

$$\Phi\big(\alpha + \beta\Phi^{-1}(\pi_{t-1}) + v_{t-1} + U_t\big) \sim \mathrm{Vas}(\pi^*, \varrho^*),$$

where the parameters π^* and ϱ^* are the solution of the equations

$$\alpha + \beta\Phi^{-1}(\pi_{t-1}) + v_{t-1} = \frac{\Phi^{-1}(\pi^*)}{\sqrt{1-\varrho^*}} \qquad \text{and} \qquad \sigma_U^2 = \frac{\varrho^*}{1-\varrho^*},$$

which is given by

$$\pi^* = p_t = \Phi\left(\frac{\alpha + \beta\Phi^{-1}(\pi_{t-1}) + v_{t-1}}{\sqrt{1+\sigma_U^2}}\right) \qquad \text{and} \qquad \varrho^* = \frac{\sigma_U^2}{1+\sigma_U^2}.$$

Now from section 8.1 follows (13) and (14).

Using the product moment

$$\begin{aligned} E[A_{t,i}A_{t,j}|\tilde{\pi}_{t-1}, V_{t-1}] &= E[E[A_{t,i}A_{t,j}|\tilde{\pi}_t, \tilde{\pi}_{t-1}, V_{t-1}]|\tilde{\pi}_{t-1}, V_{t-1}] \\ &= E[E[A_{t,i}A_{t,j}|\tilde{\pi}_t,]|\tilde{\pi}_{t-1}, V_{t-1}] \\ &= E[E[A_{t,i}|\tilde{\pi}_t]E[A_{t,j}|\tilde{\pi}_t]|\tilde{\pi}_{t-1}, V_{t-1}] \\ &= E[\tilde{\pi}_t^2|\tilde{\pi}_{t-1}, V_{t-1}] \\ &= V[\tilde{\pi}_t|\tilde{\pi}_{t-1}, V_{t-1}] + (E[\tilde{\pi}_t|\tilde{\pi}_{t-1}, V_{t-1}])^2 \end{aligned} \tag{28}$$

and relation (27) the conditional covariance reaches

$$\begin{aligned} Cov[A_{t,i}, A_{t,j}|\tilde{\pi}_{t-1}, V_{t-1}] &= E[A_{t,i}A_{t,j}|\tilde{\pi}_{t-1}, V_{t-1}] - E[A_{t,i}|\tilde{\pi}_{t-1}, V_{t-1}]E[A_{t,j}|\tilde{\pi}_{t-1}, V_{t-1}] \\ &= V[\tilde{\pi}_t|\tilde{\pi}_{t-1}, V_{t-1}], \end{aligned}$$

which is equivalent to equation (15).

Under the consideration of equations (3) and (28), it holds true that

$$V[\tilde{\pi}_t|\tilde{\pi}_{t-1} = \pi_{t-1}, V_{t-1} = v_{t-1}] = V\left[\Phi(\alpha + \beta\Phi^{-1}(\pi_{t-1}) + v_{t-1} + U_t)\right],$$

where $\Phi(\alpha + \beta\Phi^{-1}(\pi_{t-1}) + v_{t-1} + U_t)$ is Vasicek distributed with expectation (14) and variance (16).

8.5 Proof of the equivalence to the Basel II model with autoregressive systematic factor

The probit-AR(1)-model without macroeconomic impact variables

$$\begin{aligned}\Phi^{-1}(\tilde{\pi}_t) &= \alpha + \beta\,\Phi^{-1}(\tilde{\pi}_{t-1}) + U_t,\\ U_t &\overset{i.i.d.}{\sim} \mathrm{N}(0,\sigma_U^2)\end{aligned}$$

is parameterized by α, β and σ_U^2. These parameters determine

$$P[A_{t,i}=1] = E[\tilde{\pi}_t] = \Phi\left(\frac{\alpha\sqrt{1-\beta^2}}{(1-\beta)\sqrt{1-\beta^2+\sigma_U^2}}\right) =: \pi,$$

as well as

$$Cov[A_{t,i}, A_{t,j}] = V[\tilde{\pi}_t] = \Phi_2\left(\Phi^{-1}(\pi), \Phi^{-1}(\pi); \varrho\right) - \pi^2,$$

where

$$\varrho := \frac{\sigma_U^2}{1-\beta^2+\sigma_U^2}.$$

Alternatively, one can look at the Basel II model with an autoregressive systematic factor. The variables and parameters carry an asterisk as superscript in the following. The default indicators are

$$A^*_{t,i} = \mathbf{1}_{\{B^*_{t,i} < \Phi^{-1}(\pi^*)\}}$$

where the default triggering wealth variable of obligor i in period t is modeled by

$$B^*_{t,i} = \sqrt{\varrho^*}Z^*_t + \sqrt{1-\varrho^*}W^*_{t,i}$$

with

$$W^*_{t,i} \overset{i.i.d.}{\sim} \mathrm{N}(0,1).$$

The systematic factor Z^*_t is assumed to be stochastically independent of the idiosyncratic factors $W^*_{t,i}$ and to follow a stationary first-order autoregressive process

$$Z^*_t = r^* Z^*_{t-1} + M^*_t$$

with $-1 < r^* < 1$ and irregular components

$$M^*_t \overset{i.i.d.}{\sim} \mathrm{N}(0, 1-r^{*2}),$$

which leads to

$$Z^*_t \overset{i.i.d.}{\sim} \mathrm{N}(0,1).$$

This model is parameterized by π^*, ϱ^* and r^*. The unconditional default probability is given by

$$P[A^*_{t,i}=1] = \pi^*$$

and the stochastic default probability in period t can be defined as

$$\tilde{\pi}^*_t := P[A^*_{t,i}=1|Z^*_t] = \Phi\left(\frac{\Phi^{-1}(\pi^*) - \sqrt{\varrho^*}Z^*_t}{\sqrt{1-\varrho^*}}\right).$$

Therefore, the *probits* in this model take the form

$$\Phi^{-1}(\tilde{\pi}^*_t) = \frac{\Phi^{-1}(\pi^*) - \sqrt{\varrho^*}Z^*_t}{\sqrt{1-\varrho^*}},$$

which implies that

$$
\begin{aligned}
\Phi^{-1}(\tilde{\pi}_t^*) - r^*\Phi^{-1}(\tilde{\pi}_{t-1}^*) &= \frac{1-r^*}{\sqrt{1-\varrho^*}}\Phi^{-1}(\pi^*) - \sqrt{\frac{\varrho^*}{1-\varrho^*}}(Z_t^* - r^* Z_{t-1}^*) \\
&= \frac{1-r^*}{\sqrt{1-\varrho^*}}\Phi^{-1}(\pi^*) - \sqrt{\frac{\varrho^*}{1-\varrho^*}} M_t^*
\end{aligned}
$$

so that the probits have the following autoregressive representation

$$
\begin{aligned}
\Phi^{-1}(\tilde{\pi}_t^*) &= \frac{1-r^*}{\sqrt{1-\varrho^*}}\Phi^{-1}(\pi^*) + r^*\Phi^{-1}(\tilde{\pi}_{t-1}^*) - \sqrt{\frac{\varrho^*}{1-\varrho^*}} M_t^* \\
&= \alpha^* + \beta^*\Phi^{-1}(\tilde{\pi}_{t-1}^*) + U_t^*
\end{aligned}
$$

with

$$
\alpha^* := \frac{1-r^*}{\sqrt{1-\varrho^*}}\Phi^{-1}(\pi^*), \qquad \beta^* := r^* \quad \text{and} \quad U_t^* := -\sqrt{\frac{\varrho^*}{1-\varrho^*}} M_t^*.
$$

Therefore, the Basel II model with autoregressive systematic factor is nothing else than a probit-AR(1)-model, where the model parameters are related to each other as follows:

$$
\begin{aligned}
\alpha &= \frac{1-r^*}{\sqrt{1-\varrho^*}}\Phi^{-1}(\pi^*) & \pi^* &= \Phi\left(\frac{\alpha\sqrt{1-\beta^2}}{(1-\beta)\sqrt{1-\beta^2+\sigma_U^2}}\right) \\
\beta &= r^* & \varrho^* &= \frac{\sigma_U^2}{1-\beta^2+\sigma_U^2} \\
\sigma_U^2 &= \frac{\varrho^*}{1-\varrho^*}(1-r^{*2}) & r^* &= \beta.
\end{aligned}
$$

References

O. D. Anderson. *Time Series Analysis and Forecasting.* Butterworth & Co, London, 1976.

Basel Committee on Banking Supervision. International convergence of capital measurement and capital standards. http://www.bis.org/publ/bcbs107.htm, June 2004.

C. Bluhm, L. Overbeck, and C. Wagner. *An Introduction to Credit Risk.* Chapman & Hall/CRC, Boca Raton, 2003.

R. Frey and A. J. McNeil. Dependent defaults in models of portfolio credit risk. *Journal of Risk*, 6(1):59–92, 2003.

S. Höse and K. Vogl. Modeling and estimating the credit cycle by a probit-AR(1)-process. *Dresdner Beiträge zu Quantitativen Verfahren*, 44/05, 2005.

H. Joe. *Multivariate Models and Dependence Concepts.* Chapman & Hall, London, 1997.

L. Ljung. *System Identification.* Prentice Hall PTR, Upper Saddle River, 1999.

Dresdner Beiträge zu Quantitativen Verfahren (ISSN 0945-4802)

Ältere Ausgaben (1994 – 27/99): www.tu-dresden.de/wwqvs/VaR/qvreihe.htm

28/99 **S. Huschens:** Anmerkungen zur Value-at-Risk-Definition.
Erschienen in: *Datamining und Computational Finance* (Ergebnisse des 7. Karlsruher Ökonometrie-Workshops), Hrsg.: G. Bol, G. Nakhaeizadeh, K.-H. Vollmer, Physica-Verlag, Heidelberg 1999, S. 29-41.

29/99 **S. Huschens:** Verfahren zur Value-at-Risk-Berechnung im Marktrisikobereich.
Erschienen in: *Handbuch Risikomanagement,* Hrsg.: L. Johanning, B. Rudolph, Uhlenbruch-Verlag, München 2000, S. 181-218.

30/00 **S. Huschens:** Value-at-Risk-Berechnung durch historische Simulation.

31/00 **S. Huschens:** Von der Markt- zur Kreditrisikomessung.
Erschienen in: *Wertorientiertes Risiko-Management für Industrie und Handel - Methoden, Fallbeispiele, Checklisten*, Hrsg: W. Gleißner, G. Meier, Gabler-Verlag, Wiesbaden 2001, S. 389-408.

32/02 **S. Höse, S. Huschens:** Sind interne Ratingsysteme im Rahmen von Basel II evaluierbar? Zur Schätzung von Ausfallwahrscheinlichkeiten durch Ausfallquoten
Erschienen in: *Zeitschrift für Betriebswirtschaft,* 2003, Jg. 73, Heft 2, S. 139-168.

33/03 **S. Höse, S. Huschens:** From Credit Scores to Stable Default Probabilities: A Model Based Approach.
Erschienen in: *Exploratory Data Analysis in Empirical Research,* Hrsg.: M. Schwaiger, O. Opitz, Springer, Berlin, 2003, S. 454-462.

34/03 **S. Höse, S. Huschens:** Estimation of Default Probabilities in a Single-Factor Model.
Erschienen in: *Between Data Science and Applied Data Analysis,* Hrsg.: M. Schrader, W. Gaul, M. Vichi, Springer, Berlin, 2003, S. 546-554.

35/03 **S. Höse, S. Huschens:** Simultaneous Confidence Intervals for Default Probabilities.
Erschienen in: *Between Data Science and Applied Data Analysis,* Hrsg.: M. Schrader, W. Gaul, M. Vichi, Springer, Berlin, 2003, S. 555-560.

36/03 **S. Huschens, K. Vogl, R. Wania:** Estimation of Default Probabilities and Default Correlations.
Erschienen in: *Risk Management*, Hrsg: M. Frenkel, U. Hommel, M. Rudolf, Springer, 2. Auflage, Berlin, 2005, S. 239-259.

37/04 **K. Vogl, R. Wania:** BLUEs for Default Probabilities.

38/04 **A. Henking:** Simultane Validierung von Ausfallwahrscheinlichkeiten.

39/04 **S. Huschens, G. Stahl:** A General Framework for IRBA Backtesting.
Erschienen in: *Bankarchiv, Zeitschrift für das gesamte Bank- und Börsenwesen*, Hrsg: Österreichische bankwissenschaftliche Gesellschaft, Springer, Wien, April 2005, Jg. 53, S. 241-248.

40/04 **S. Huschens:** Backtesting von Ausfallwahrscheinlichkeiten.
Erscheint 2005.

41/04 **S. Huschens:** Dreizehn Korrelationen in Kreditrisikomodellen.
Erschienen in: *Banken, Finanzierung und Unternehmensführung*, Hrsg: T. Burkhardt, Duncker & Humblot, Berlin, 2004, S. 177-188.

42/04 **S. Huschens:** Faktorstruktur und Marktmodelle.
Erscheint 2005.

43/04 **S. Huschens, G. Stahl:** Granularität dominiert Korrelation.
Erschienen in: *RiskNews*, Vol. 1, Heft 6, 2004, S. 28-29.

44/05 **S. Höse, K. Vogl:** Modeling and Estimating the Credit Cycle by a Probit-AR(1)-Process.

45/05 **S. Höse, K. Vogl:** Predicting the Credit Cycle with an Autoregressive Model.