

To appear in: [International Journal of Forecasting](#)

An Empirical Comparison of Default Risk Forecasts from Alternative Credit Rating Philosophies

Daniel Rösch*

University of Regensburg

* *Dr. Daniel Rösch*, Department of Statistics, Faculty of Business Management, Economics and Management Information Systems, University of Regensburg, 93040 Regensburg, Germany
Phone: +49-941-943-2752, Fax : +49-941-943-4936
Email: Daniel.Roesch@wiwi.uni-regensburg.de,
Internet: <http://www.danielroesch.de>

I gratefully acknowledge helpful suggestions from two anonymous referees. I also thank participants at the 2nd International Conference on Credit Risk, April 15-16 2004 in Montreal for helpful comments.

Abstract

The New Basel Capital Accord will allow the determination of banks' regulatory capital requirements due to probabilities of default which are estimated and forecasted from internal ratings. Broadly, two rating philosophies are distinguished: Through the Cycle versus Point in Time Ratings. We employ a Likelihood-Ratio backtesting of both types with respect to their probability of default forecasts and correlations derived from a nonlinear random effects panel model using data from Standard & Poor's. The implications for risk capital using these different philosophies are demonstrated. It is shown that Point in Time Ratings will exhibit much lower correlations and, thus, default probability forecasts should be more precise. As a consequence, Value at Risk quantiles of default distributions should be lower than those generated by Through the Cycle Ratings. Nevertheless banks which use Point in Time Ratings may be punished in times of economic stress if the implied reduction of asset correlation is not taken into account.

Key Words:

Credit Rating, Basel II, Backtesting, Risk Management, Credit Risk Modeling

1 The Problem

The planned Basel Accord for the revision of minimum requirements for banks' risk capital has raised a lot of discussions about how to measure credit risk and forecast probabilities of default (PDs)¹. Within the new revisions, aimed to take effect by the end of 2006, banks are allowed to determine their capital charges due to the inherent credit risk of each borrower. This credit risk, or the probabilities of default respectively, can be inferred to a bank from an internal credit rating model. The planned approach is therefore called the "Internal Ratings Based" approach. A further driver for regulatory capital is the correlation between borrowers. However, this parameter is pre-specified by the supervising authorities. A bank's internal estimates of correlations are not expected to be used for capital charges.

Usually one distinguishes two types of credit rating philosophies, see e.g. Basel Committee on Banking Supervision (2000a,b): Through the Cycle versus Point in Time Ratings. The first group is mainly employed by external credit rating agencies such as Moody's and Standard & Poor's (S&P) while most banks internally follow the second philosophy, see Treacy and Carey (2000). Each philosophy has its own characteristics and purposes. Rating agencies focus on the long term over one or more business cycles. That is, they provide ratings which are forward-looking and do not try to offer a snapshot of the present situation or the near future, see Standard & Poor's (2002). A similar interpretation by Moody's can be found in Catarineu-Rabell, Jackson and Tsomocos

¹ See Basel Committee on Banking Supervision (2003, 2002, 2001).

(2003). As such, an assigned rating is nearly constant over time (Basel Committee on Banking Supervision, 2000a) and is not conditioned on the point of the cycle (Catarineu-Rabell, Jackson and Tsomocos, 2003).

Borrowers are grouped into rating grades which are abbreviated with letters and/or ciphers. For example, S&P use grades from “AAA” (“Highest Rating; The obligor’s capacity to meet its financial commitment on the obligation is extremely strong”) over “AA”, “A”, “BBB”, and so on, to “C” (“A bankruptcy petition has been filed or similar action has been taken but payments on this obligation are being continued”). Default probabilities are assigned to a grade by calculating the observed default rate of all borrowers within this grade in each year and averaging these figures over a historical horizon (Standard & Poor’s, 2001).

A Point in Time Rating on the other hand reflects a borrower’s situation and the most likely future condition over an exactly pre-specified horizon, e.g. one year (Basel Committee on Banking Supervision, 2000a). Therefore, the rating changes as soon as the borrower’s condition changes within a business cycle and, thus, the ratings are more volatile than the Through the Cycle Ratings (Catarineu-Rabell, Jackson and Tsomocos, 2003, see also Carey and Hrycay, 2001). A well-established paradigm of a Point in

Time Rating is the proprietary Merton style model from KMV² which makes use of current equity price information, see Crosbie (1998) for an overview.

As such, a Point in Time Rating incorporates all relevant information which influences the one-year creditworthiness of a borrower, i.e. the probability that the borrower will default within the next year. Point in Time Ratings and default probabilities are usually derived from market data (e.g. equity returns, credit spreads) as in the KMV model or from statistical models, such as discriminant analysis or logistic regression. These types of ratings are often used to calculate *economic* capital.

Within the proposals of the new Basel Accord, there is no explicit guidance on which type of rating philosophy should be employed for the calculation of *regulatory* capital requirements although the philosophies, or respectively the default probabilities which they generate, are essential for the new capital adequacy framework. Thus, researchers have begun to analyse rating philosophies empirically. Carey and Hrycay (2001) analyse the effects of calibrating external ratings to banks. Crouhy, Galai and Mark (2001) argue that Point in Time Ratings are more appropriate for the purposes of capital allocation. Catarineu-Rabell, Jackson and Tsomocos (2003) suggest that using Through the Cycle Ratings may mitigate the problem of procyclical capital requirements.

² KMV is a subsidiary of Moody's. The three letters are the initials of their founders, Steven Kealhofer, John McQuown and Oldrich Vasicek.

To summarize, the delineation above shows that an exact definition of a Point in Time Rating is possible – it reflects a borrower’s one-year probability of default -, while a definition of a Through the Cycle Rating is not as clear-cut. Regarding the information content, there is some evidence that Through the Cycle Ratings do not fully reflect all available information, see Altman and Kao (1992), Lando and Skodeberg (2002), and the comments in Löffler (2004).

In the context of the discussion on rating philosophies, the present paper tries to make several contributions. Firstly, we compare default probability estimates and estimates for asset correlation for Through the Cycle and Point in Time Ratings. Default data from S&P are used and we show that correlations implied by Through the Cycle ratings are merely substitutes for fluctuating underlying default probabilities of the rating grades over time. Using S&P’s Through the Cycle Rating as a starting point we generate a “mimicking” Point in Time Rating by adding information about the state of the business cycle. It is shown that asset correlations using this mimicking Point in Time Rating are much smaller than in the Through the Cycle case since one-year default probabilities are reflected more adequately.

Secondly, we analyse which rating philosophy is better in forecasting defaults. We do this by employing the Likelihood Ratio test as it is suggested in Berkowitz (2001) in the context of market risk. Although the test results do not completely answer the question which rating scheme should be favoured, lastly we assess some implications of both types of rating schemes for economic and regulatory capital.

The rest of the paper is organized as follows. Section 2 describes the estimation approach for default probabilities and asset correlations and the backtesting procedure. Section 3 contains the empirical results of the estimations, the backtesting and the implications for capital requirements. Section 4 concludes.

2 Model and Backtesting Approach

2.1 Generating Forecasts for Default Distributions

Forecasts of the distributions of potential default events occurring in a future period are compared for two kinds of rating philosophies: Through the Cycle and Point in Time Ratings. While default data for Through the Cycle Ratings are publicly available, for example by S&P, the data for Point in Time Ratings frequently employed by banks are proprietary. Even if point in time default data were available, a problem would arise in comparing them since the databases used could be very different. In particular, banks often have loans outstanding to small and medium sized borrowers while S&P rate issuers of traded bonds.

To overcome this problem we use a different approach. Estimates for default probabilities generated from Through the Cycle Ratings are directly calculated by S&P's one year default rates. In order to achieve default probability estimates for Point in Time Ratings of the same database we "mimic" Point in Time Ratings by including further

information about the point of the business cycle into the rating. S&P's ratings are well known to reflect views about borrowers over the whole credit cycle. Thus, by including additional time-dependent information about the actual state of the economy a Point in Time Rating, or its default probability estimates respectively, can be proxied. The procedure used here is similar to the one used by Hamerle, Liebig and Rösch (2003a), Rösch (2003), and Hamerle, Liebig and Rösch (2003b) and is described below.

As in Gordy (2003) it is assumed that borrowers can be grouped into homogenous risk buckets, for example rating grades or industries. A borrower i within a bucket defaults if the discrete time return S_{it} on the firm's assets falls below some threshold c at time t ($i=1, \dots, N_t, t=1, \dots, T$). In finance it is frequently assumed that the return is modelled as a normal distributed random variate $S_{it} \sim N(\delta, \gamma)$, where δ and γ denote mean and standard deviation respectively. As under the Basel II Internal Ratings Based approach the asset return is treated as a latent variable and the one factor model

$$S_{it} = \delta + \tilde{b} F_t + \varpi U_{it}$$

is assumed where

$$F_t \sim N(0,1), \quad U_{it} \sim N(0,1)$$

$(i=1, \dots, N_t, t=1, \dots, T)$ are normally distributed with mean zero and standard deviation one, and \tilde{b} and ϖ are parameters. Idiosyncratic shocks U_{it} are assumed to be independent from the systematic factor F_t and independent for different borrowers. All random variables are serially independent. It holds that $\gamma^2 = \tilde{b}^2 + \varpi^2$. Alternatively, the normalized return $R_{it} = \frac{S_{it} - \delta}{\gamma}$ on the firm's assets is given by

$$R_{it} = b F_t + \sqrt{1 - b^2} U_{it} \quad (1)$$

where $b = \tilde{b}/\gamma$ denotes the exposure to the common factor.

Under these assumptions the correlation between the normalized asset returns of any two borrowers is b^2 . A detailed description and implications of the model can be found in Gordy (2003). In the Basel II proposal as of January 2001 this correlation is set to 20%. In the proposals as of October 2002 and April 2003 the correlation decreases with increasing default probability from 24% to 12%.

The probability of default at time t for borrower i within a given bucket, given survival before t , is

$$\begin{aligned}
\lambda &= P(S_{it} < c) = P\left(R_{it} < \frac{c - \delta}{\gamma}\right) \\
&= P(R_{it} < \beta_0) = P\left(b F_t + \sqrt{1 - b^2} U_{it} < \beta_0\right) = \Phi(\beta_0)
\end{aligned} \tag{2}$$

where $\beta_0 = (c - \delta)/\gamma$ and $\Phi(\cdot)$ denotes the cumulative standard normal distribution function. Conditional on a realization f_t of the common factor the probability of default is

$$\lambda(f_t) = P(R_{it} < \beta_0 | f_t) = P\left(U_{it} < \frac{\beta_0 - b f_t}{\sqrt{1 - b^2}}\right) = \Phi\left(\frac{\beta_0 - b f_t}{\sqrt{1 - b^2}}\right). \tag{3}$$

Model (2) assumes that there is a default threshold which is time invariant and, thus, that there is a default probability which is constant over time. However, if a rating agency attempts to rate through the cycle the actual one year default probability may be different since the information about the current point of the business cycle is intentionally not taken into account. Rather some kind of average default probability over the cycle is reflected.

In order to mimic default probabilities generated by a point in time rating, the default threshold and, thus the default probability, is modeled as a function of the business cycle, for which we use macroeconomic key figures (i.e. risk factors) as proxies. Let $z_t = (z_{1t}, \dots, z_{Kt})'$ denote a K -vector of risk factors and $\tilde{\beta} = (\tilde{\beta}_1, \dots, \tilde{\beta}_K)'$ the vector of

sensitivities with regard to these factors. Then, within a bucket, the mean of the asset returns is modeled as a linear function of the observable factors, i.e. $E(S_{it}) = \delta_t = \delta + \tilde{\boldsymbol{\beta}}' \mathbf{z}_t$ and therefore is assumed to depend on time. In an economic boom the z-factors boost the expected return while it is decreased in a recession. The asset return can then be written as dependent on the observable factors as

$$S_{it} = \delta + \tilde{\boldsymbol{\beta}}' \mathbf{z}_t + \tilde{b} F_t + \varpi U_{it}$$

and the default probability conditional on the observable risk factors is given by

$$\begin{aligned} \lambda(\mathbf{z}_t) &= P(S_{it} < c) = P\left(\frac{S_{it} - \delta_t}{\gamma} < \frac{c - \delta_t}{\gamma}\right) = P(R_{it} < \beta_0 + \boldsymbol{\beta}' \mathbf{z}_t) \\ &= P\left(b F_t + \sqrt{1 - b^2} U_{it} < \beta_0 + \boldsymbol{\beta}' \mathbf{z}_t\right) = \Phi(\beta_0 + \boldsymbol{\beta}' \mathbf{z}_t) \end{aligned} \quad (4)$$

where $\beta_0 = (c - \delta)/\gamma$ and $\boldsymbol{\beta} = -\tilde{\boldsymbol{\beta}}/\gamma$ denote the transformed threshold and vector of factor exposures. Conditional on a realization f_t of the common factor the probability of default is

$$\lambda(\mathbf{z}_t, f_t) = P\left(U_{it} < \frac{\beta_0 + \boldsymbol{\beta}' \mathbf{z}_t - b f_t}{\sqrt{1 - b^2}}\right) = \Phi\left(\frac{\beta_0 + \boldsymbol{\beta}' \mathbf{z}_t - b f_t}{\sqrt{1 - b^2}}\right). \quad (5)$$

An advantage of models (3) and (5) is that neither need to observe returns on the firms' assets themselves. Instead, a time series of numbers of obligors and observed default events is sufficient for parameter estimation. Suppose for a given segment (for example a rating grade or an industry sector) one observes a time series of defaults d_t , and numbers N_t of borrowers ($t=1, \dots, T$). From the independence of the idiosyncratic shocks U_{it} between obligors it follows that the default events are independent conditional on the realization f_t of the common random factor. Thus, the number of defaults is a conditional binomial with parameters N_t and default probability $\lambda(f_t)$, or $\lambda(z_t, f_t)$. Furthermore, all random variables are assumed to be serially independent. Then the log-likelihood function for the observed time series is set up by integrating the conditional binomial distribution over the standard normal distributed random effect and summing the logarithms of these unconditional distributions over the T years, i.e. for example for model (3)

$$l(\beta_0, b) = \sum_{t=1}^T \ln \left\{ \int_{-\infty}^{\infty} \binom{N_t}{d_t} \lambda(f_t)^{d_t} \cdot [1 - \lambda(f_t)]^{(N_t - d_t)} \varphi(f_t) df_t \right\}$$

where $\varphi(\cdot)$ denotes the density function of the standard normal distribution. Given the time series of default data the parameters of (3) and (5) can be estimated by Maximum Likelihood (see Gordy and Heitfield, 2000, Hamerle, Liebig and Rösch, 2003a and Rösch, 2003).

It is important to note that for each of the rating schemes, given estimates for default probabilities and correlations, the default distributions can be forecasted, i.e. the distribution of the potential numbers of defaulting companies for the next period $T+1$ (e.g. one year) can be calculated as shown in Vasicek (1987). The probability distribution for the number D_{T+1} of defaulting companies within a risk segment, given the number N_{T+1} of companies in this segment at the beginning of the period is

$$g(d_{T+1}) = \begin{cases} \binom{N_{T+1}}{d_{T+1}} \cdot \int_{-\infty}^{+\infty} [\lambda(f_{T+1})^{d_{T+1}} \cdot [1 - \lambda(f_{T+1})]^{(N_{T+1} - d_{T+1})}] \varphi(f_{T+1}) df_{T+1} & d_{T+1} = 0, 1, 2, \dots, N_{T+1} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where $\varphi(\cdot)$ denotes the density function of the standard normal distribution and $\lambda(f_{T+1})$ denotes the conditional default probability in (3) or (5) depending on which model is assumed. Due to the serial independence of the random factor the defaults are also serially independent. From the number of defaults the default rate Q_{T+1} can easily be calculated as $Q_{T+1} = \frac{D_{T+1}}{N_{T+1}}$. Furthermore, Vasicek (1991) extends the distribution

for the default rate to the limiting case of a portfolio with infinitely many obligors.

2.2 Backtesting Credit Risk Models

Many applications in forecasting focus on point forecasting. In the context of financial risk management however, most models generate outputs which cannot be summarized by a single scalar. For example in market risk, one is interested in the whole distribution of potential market values of a portfolio within a pre-specified holding period, e.g. one or ten trading days. Frequently, the information about the whole distribution is condensed into interval forecasts, such as the Value-at-Risk concept, which measures the portfolio loss which will not be exceeded within the holding period with a given probability. Since 1997, these forecasts generated by so-called “internal models” are allowed to be used by banks to determine regulatory capital requirements for market risk. The validity of the forecasts is then assessed by supervisors who compare the model outcomes with actual losses. They follow a simple backtesting rule which is based on the likelihood of occurrences of exceeding the forecasted Value-at-Risk within 250 trading days.

In applying these types of backtesting to credit risk models one faces at least two difficulties. Firstly, historical data of loan performances are typically scarce since the use of credit risk models by banks is a recent innovation. Secondly, in contrast to the changes in value of trading portfolios, defaults on loans are usually observed during longer intervals, e.g. once a year. Thus, focussing on occurrences of exceeding Value-at-Risks of default risky portfolios would require many years of observations.

In order to mitigate these problems, Berkowitz (2001) proposed a backtesting procedure which Frerichs and Löffler (2003) analysed using Monte-Carlo simulation. The approach makes use of the information about the whole distribution of potential defaults.

In general, the two rating schemes will produce different forecasts for default distributions for two reasons even if the underlying portfolio is the same: The center of the default distribution depends on the predicted default probability and its shape depends on the correlation. The empirical question which should be answered then is: Which distribution is more adequate for forecasting?

The backtesting approach works as follows. Suppose we have estimated model (3) and obtained estimates $\hat{\beta}_0$ and \hat{b} for the relevant parameters β_0 and b . Then, using these estimates we can forecast the probability distribution of potential default for period $T + 1$ analogously to (6). Denoting the forecast for the conditional probability of default which makes use of the parameter estimates with $\hat{\lambda}(f_{T+1}) = \lambda(f_{T+1}, \hat{\beta}_0, \hat{b})$, one obtains

$$\hat{g}(d_{T+1}) = \begin{cases} \binom{N_{T+1}}{d_{T+1}} \cdot \int_{-\infty}^{+\infty} \left[\hat{\lambda}(f_{T+1})^{d_{T+1}} \cdot [1 - \hat{\lambda}(f_{T+1})]^{(N_{T+1} - d_{T+1})} \right] \varphi(f_{T+1}) df_{T+1} & d_{T+1} = 0, 1, 2, \dots, N_{T+1} \\ 0 & \text{otherwise} \end{cases}$$

To explain the main idea of the backtesting approach, let us assume for a moment that the number of obligors is infinitely large. Then the density of the default rate Q_{T+1} in the next year is given by the formula derived by Vasicek (1991) or Koyluoglu and

Hickman (1998a,b). Let $\hat{g}(q_{T+1})$ denote the forecasted density for the default rate.

Within the next year the companies' default indicators will be observable and at the end of the year we can observe the (ex post) default rate q_{T+1} which serves as a benchmark for backtesting. Then, as Rosenblatt (1952) shows, the transformation

$$x_{T+1} = \int_{-\infty}^{q_{T+1}} \hat{g}(u) du$$

is i.i.d. uniformly distributed on (0,1). Furthermore, Berkowitz (2001) shows that the transformation

$$z_{T+1} = \Phi^{-1}(x_{T+1}) = \Phi^{-1}\left(\int_{-\infty}^{q_{T+1}} \hat{g}(u) du\right) \quad (7)$$

is an i.i.d. standard normal variate if the forecasted probability density is correct.

For a sequence of L forecasted distributions and serially uncorrelated ex post observed default rates the transformation (7) can be conducted and a test statistic based on the likelihood ratio under the observations and under the null hypothesis can be calculated as

$$\begin{aligned}
LR &= -2 \cdot (L(0,1) - L(\hat{\mu}, \hat{\sigma})) = \\
&= -2 \cdot \left(\left(-\frac{L}{2} \log 2\pi - \sum_{l=1}^L \frac{z_{T+l}^2}{2} \right) - \left(-\frac{L}{2} \log 2\pi - \frac{L}{2} \log \hat{\sigma}^2 - \sum_{l=1}^L \frac{(z_{T+l} - \hat{\mu})^2}{2\hat{\sigma}^2} \right) \right)
\end{aligned} \tag{8}$$

where $\hat{\mu} = \frac{1}{L} \sum_{l=1}^L z_{T+l}$ and $\hat{\sigma}^2 = \frac{1}{L} \sum_{l=1}^L (z_{T+l} - \hat{\mu})^2$ are the Maximum Likelihood estimators for the mean and the variance of the transformed variable. Under the null hypothesis, the test statistic is asymptotically $\chi^2(2)$ -distributed. Note that this outline refers to the case of a portfolio with infinitely many obligors. In reality, this will not be the case, and the number of obligors will be finite, but sufficiently large. Thus, the above results will hold only approximately. However, Frerichs and Löffler (2003) show that the backtesting approach also works well with a finite though large number of borrowers.

3 Results

3.1 The Data and Estimation Results

We use Standard & Poor's (2001) transition matrices from 1982 to 2000 which are publicly available. Firstly model (3) is estimated for each rating grade. Actually the ML-estimates could only be established for the so-called "speculative grades"

- BB (“Obligation faces major ongoing uncertainties or exposure to adverse business, financial, or economic conditions”),
- B (“Adverse conditions will likely impair the obligor’s capacity or willingness to meet its financial commitment”) and
- CCC (“Currently vulnerable to non-payment”)

since the numbers of defaults in the upper grades are too small. We estimate a constant default probability over the whole cycle, and an asset correlation which models the fluctuations around this Through the Cycle default probability. Table 1 contains the results.

[---Insert Table 1 about here---]

As it can be seen, the constant which represents the normalized default threshold increases with decreasing rating grade. Furthermore, the estimate for the asset correlation which is the squared coefficient b^2 is about 5.3% for grade BB, 4.4% for grade B, and 6.6% for grade CCC. These values are much lower than those assumed in the Basel II Accord and their magnitudes correspond to those in Gordy (2000).

In the next step model (5) is estimated with lagged macroeconomic variables as additional risk factors which were provided by the OECD and Deutsche Bundesbank. By adding proxies for the point of the business cycle we try to mimic a Point in Time Rating scheme. Table 2 contains the results of the estimation.

[---Insert Table 2 about here---]

In grades BB and CCC one variable was enough to “explain” the correlation and render the coefficient of the random effect insignificant. In grade BB, an increase of the Federal Funds Rate comes along with an increase of default probabilities in the following year. This is reasonable because higher rates may lead to higher interest rates for debt. In grade CCC, higher unemployment is associated with decreasing default probabilities in the next year. This is also plausible for two reasons. Firstly, if firms rationalise they release employees. In the following years, their cost pressures decrease leading to lower default risk. Secondly, the government may stimulate the economy by higher public expenditure in times of higher unemployment. This could also decrease default risk. In grade B, although four variables were included, the random effect coefficient is still significantly different from zero but the remaining asset correlation could be reduced to less than 1%.

3.2 *Backtesting of Forecasted Distributions*

So far, we have two rating schemes and try to accomplish a backtesting for the adequacy of the default distribution forecasts they produce. In addition, we compare the ratings to a “naïve” forecast procedure which uses the realized default rate in a given year as forecast for next year’s default probability. Since in Basel II the minimum re-

quirements envisage at least five years of historical data we use five years, the period from 1996 to 2000.

For the Through the Cycle Rating we proceed as follows. From the history of defaults until 1995 we estimate default probabilities by averaging the default rates of each rating grade. This average is used as forecast for the default probability of 1996. The asset correlation is determined according to the Basel proposal as of October 2002 depending on the default probability forecast. Then we average default rates until 1996 and use this average as forecast for the default probability of 1997, and so on. These results are then compared to the actual numbers of defaults in the corresponding year which can be extracted from the number of issuers and the default rates in the transition matrices of Standard & Poor's (2001). In summary we obtain five years of default probability forecasts, asset correlations, issuers in the rating grade and, in addition, realized defaults in the corresponding year. The "naïve" approach works similarly but uses only last year's default rate as forecast instead of the average over the entire history. The results are summarized in Table 3 for the "naïve" alternative and in Table 4 for the Through the Cycle scheme.

[---Insert Tables 3 and 4 about here---]

In the case of the Point in Time Rating we firstly estimate the models (5) for the rating grades with the data until 1995. Since our macroeconomic factors exhibit time lags of one and two years, we substitute the realized values of 1994 and 1995 into the estima-

tion equation and get forecasts for the default probabilities and the asset correlations due to the model for year 1996. Then the model is estimated with the data until 1996 and the default probabilities and asset correlations are estimated for 1997 the same way. We repeat this proceeding up to the forecasts for year 2000. In summary, we also get five years of model default probability forecasts and model asset correlations which are summarized in Table 5 for each of the three rating grades. Figures 1 to 3 show the actual default rates of the three rating grades and the corresponding default probability estimates due to both rating schemes. As it can be seen, the Point in Time Rating default probabilities seem to result in a more adequate fit to the default rates, in particular for grade B.

[---Insert Table 5 and Exhibits 1 to 3 about here---]

Given these three sets of default probability forecasts, asset correlations and issuers for each grade, distributions of potential defaults can be forecasted according to (6) for each year and rating scheme. Then, using actual defaults the Likelihood Ratios can be computed using the five year history for each rating. Note that the number of issuers in grade CCC is small. The outcomes should thus be treated with some caution. The results of the LR test are summarized in Table 6.

[---Insert Table 6 about here---]

The “naïve” procedure can be rejected for grades BB and B, signalling that a single year of observed default rates is not adequate for forecasting. Regarding the more advanced approaches the tests indicate that the predicted distributions are not incorrect for each rating scheme and grade except for grade BB in the case of the Through the Cycle Rating. Here the null hypothesis can be rejected at the 10% level. Thus, at least two questions arise. The first is about the power of the test, in particular for the short time series of only five years each. The second is about the implications of these test results. If both rating schemes cannot be rejected, which one should be favoured for calculating capital charges? We address the first question in the next section and the second question in section 3.4.

3.3 Power Considerations of the LR Test

Most of the power considerations of the LR test can be found in Frerichs and Löffler (2003). They show that the LR test performs in a satisfactory manner even for small time series. Here we amend their results for the present case. We analyse the power of the test for grades BB and B which exhibit default probabilities of approximately 0.01 and 0.05 respectively. Therefore, as in Frerichs and Löffler (2003) we generate portfolios of homogeneous borrowers with 0.01 and 0.05 default probability each and ask if the test detects deviations from the null hypothesis. Since the main driver for portfolio losses is the asset correlation we only alter this parameter. 10,000 simulations are run in each setting. Table 7 summarizes the rejections of the null hypothesis for grade BB with

a default probability of 0.01, portfolio sizes of 1,000 and 10,000 and different assumptions about asset correlations.

[---Insert Tables 7 and 8 about here---]

In the first three rows a true correlation of 0.02 is assumed. This hypothesis is falsely rejected in about 16.9% of all cases if it is true for the portfolio of 1,000 borrowers and in about 16% of all cases for the portfolio of 10,000 borrowers at the 10% significance level. Thus, the significance level is exceeded slightly too often. If the model under consideration assumes a false correlation of 0.01, as it is about the magnitude in our empirical case the power is low. For the 1,000 borrower portfolio the null is rejected in only 18% of all cases and in the 10,000 borrower case only in 29% of all cases. If the model assumes an asset correlation of 0.2 as in the Basel II proposal the power is much better. The false null is rejected in 92% and 99% respectively. Clearly, the power increases with increasing true correlation for the test with a model correlation of 0.01 and decreases for the test with a model correlation of 0.2. The significance level is always slightly exceeded if the null is true.

If the true default probability is 0.05 the test performs much better. If the true correlation is 0.02 and the correlation in the tested model is 0.002 as it is in three years of the Point of Time Rating model for grade B the null is rejected in about 64% (1,000 borrowers) and 90% (10,000 borrowers) of all cases at a significance level of 10%. If the Basel II correlation of 0.13 is used the model is rejected in 80% (1,000 borrowers) and

86% (10,000 borrowers). Again the power of the test regarding a false model with correlation of 0.002 increases with increasing true correlation and it decreases for a correlation of 0.13.

Without trying to stress the results too much the following preliminary conclusions are obvious. Empirically the test rejects only the Through the Cycle Rating of grade BB but not the Point in Time Rating. If one considers the power of the test the true correlation should be lower than 0.05. Considering grade B where both tests are not rejected the power simulations of the test indicate that the true correlation should not be much higher than about 0.02 where the rejection rates of both null hypotheses are similar. While the power considerations of the test can not completely answer the question about favouring Through the Cycle forecasts versus Point in Time forecasts, we now turn to some practical aspects of both rating schemes.

3.4 *Economic and Regulatory Implications*

For expository purposes, the distributions of potential default events are forecasted for year 2000, the last year of the backtesting horizon, and are depicted in Figures 4 to 6 for each rating grade and “advanced” rating scheme. The “naïve” alternative is no longer considered due to the rejection of its validity.

[---Insert Exhibits 4 to 6 about here---]

As can be seen from Figures 1 to 3 the year 2000 was a “bad year” with high default rates. Thus, except for grade BB the center of the default distribution due to the Point in Time Rating is shifted to the right compared with the Through the Cycle Rating. Table 9 contains the Expected Losses (EL) under both distributions.

[---Insert Table 9 about here---]

For grade B the Expected Loss is 9.8% versus 5.1%, for grade CCC 33.8% vs. 19.2%. Only for grade B the Expected Loss is 0.6% versus 1%. However, due to the lower asset correlation the default distributions are in any case narrower under the Point in Time Rating scheme. This can also be seen in Table 9 which exhibits the Value at Risk quantiles of both distribution for each grade. Moreover the Unexpected Losses (UL), defined as the distance of the quantile from the Unexpected Loss, are in any case much smaller than under the Through the Cycle Rating. That is, although 2000 is a bad year and the EL may increase, due to the more exact measurement of one-year risk the uncertainty about the distribution of potential defaults and thus the *economic* capital decreases, which is needed for buffering unexpected losses.

These findings may also have implications for *regulatory* capital requirements under the Basel II Accord. Suppose Bank A estimates default probabilities by the Through the Cycle scheme. Then the default probability estimates reflect a long run perspective of the borrowers' conditions and short term fluctuations are captured by the asset correla-

tion. The default probabilities of the rating grades will not substantially change during economic downturns and regulatory capital requirements will be stable through time.

Now consider Bank B which estimates default probabilities by a Point in Time Rating. In a recession the short term credit qualities - or the one year default probabilities - of most borrowers tend to rise. The higher default probability forecasts will be reflected in a rise of the banks' regulatory capital requirements. This is usually known as the procyclicality effect, see e.g. Allen and Saunders (2002) and the literature overview therein.

Both banks estimate default probabilities of their borrowers via internal models, but only Bank B is penalized since it must hold a higher capital buffer in the recession. The crucial point lies in the separation of default probability estimation and asset correlation under the Basel II Accord. Bank B may exhibit much lower correlations than Bank A. Nevertheless, the *presumed* asset correlation for capital requirements are the same for both and, thus, the capital requirements for Bank B may oscillate more strongly through time. A solution to the problem could be to encourage banks to use a Through the Cycle Rating, see Catarineu-Rabell, Jackson and Tsomocos (2003). However, in contrast to Point in Time Ratings which refer to one-year default probabilities, a clear definition of Through the Cycle Ratings and their default probabilities is absent. Without explicit guidance, the rating schemes may then be arbitrary. Moreover, the revision of the capital accord is meant to link regulatory capital requirements more closely to a bank's eco-

conomic capital, that is its actual one year risk, see for example Crouhy, Galai and Mark (2001).

If one agrees that the New Capital Accord aims at a more transparent credit risk buffering and that increasing default risk is natural in a recession then the cyclical behavior of capital requirements is coherent. An alternative solution may consist in an integrated view of default probabilities and correlations. Banks which demonstrate that their default probabilities are based on Point in Time may gain a reduction of the presumed correlation since their economic risk is also lower.

4 Conclusion

Some differences between forecasts generated by Through the Cycle and Point in Time Ratings were outlined empirically. Both rating philosophies should be distinguished not only by the default probability estimates which they produce but also by their implied asset correlations. While Through the Cycle default probabilities are relatively stable through time and short term fluctuations are captured by the asset correlation, Point in Time Ratings produce dynamic short term (e.g. one-year) default probabilities. Due to modeling time-varying default probabilities, asset correlations are much lower.

Both philosophies and the forecasts for default distributions may be compatible with actual default data, if correlations are correctly taken into account. For large portfolios an LR test can be employed for testing on the accuracy of the forecast distributions. The

former results in a default distribution which is wide and stable through time while the latter produces forecasts which are narrower but dynamic. Thus, default probabilities and defaults may be forecast with less noise.

Although such predictions with less uncertainty should be preferable and lead to lower economic risk, *ceteris paribus*; Point in Time Ratings may be penalized under Basel II in economic downturns when borrowers' short term default probabilities rise. This drawback stems from the presumed fixed asset correlation under Basel II which does not take into account the lower correlation and the lower economic risk which the Point in Time Rating produces.

References

- Allen, L., & Saunders, A. (2002). A Survey of Cyclical Effects in Credit Risk Measurement Models, *Working Paper*, available under www.defaultrisk.com.
- Altman, E., & Kao, D.L. (1992). Rating Drift in High Yield Bonds, *Journal of Fixed Income*, 2, 15-20.
- Basel Committee on Banking Supervision (2003). The New Basel Capital Accord, *Consultative Document*, Bank for International Settlements, April, available under www.bis.org.

- Basel Committee on Banking Supervision (2002). Quantitative Impact Study 3 Technical Guidance, Bank for International Settlements, October, available under www.bis.org.
- Basel Committee on Banking Supervision (2001). The New Basel Capital Accord, *Consultative Document*, Bank for International Settlements, January, available under www.bis.org.
- Basel Committee on Banking Supervision (2000b). Credit Ratings and Complementary Sources of Credit Quality Information, *Working Paper*, No. 3 August, available under www.bis.org.
- Basel Committee on Banking Supervision (2000a). Range of Practice on Banks' Internal Ratings Systems, *Working Paper*, January, available under www.bis.org.
- Berkowitz, J. (2001). Testing Density Forecasts, With Applications to Risk Management, *Journal of Business and Economics Statistics*, 19, 465-474.
- Carey, M., & Hrycay, M. (2001). Parameterizing Credit Risk Models With Rating Data, *Journal of Banking and Finance*, 25, 197-270.
- Catarineu-Rabell, E., Jackson, P., & Tsomocos, D.P. (2003). Procyclicality and the New Basel Accord – Banks' Choice of Loan Rating System, Working Paper No. 181, *Bank of England*, available under <http://www.bankofengland.co.uk/workingpapers/wp181.pdf>.
- Crosbie, P. (1998). Modeling Default Risk, *KMV Corporation*, San Francisco, available under www.kmv.com.

- Crouhy, M., Galai, D., & Mark, R. (2001). Prototype Risk Rating System, *Journal of Banking and Finance*, 25, 47-95.
- Frerichs, H., & Löffler, G. (2003). Evaluating Credit Risk Models Using Loss Density Forecasts, *Journal of Risk*, 5 (4), 1-23.
- Gordy, M.B. (2003). A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules, *Journal of Financial Intermediation*, 12, 199-232.
- Gordy, M.B. (2000). A Comparative Anatomy of Credit Risk Models, *Journal of Banking and Finance*, 24, 119-149.
- Gordy, M.B., & Heitfeld, N. (2000). Estimating Factor Loadings When Ratings Performance Data Are Scarce, Technical Report, *Board of Governors of the Federal Reserve System, Division of Research and Statistics*,.
- Hamerle, A., Liebig, T., & Rösch, D. (2003b). Benchmarking Asset Correlations, *Risk*, 16, November, 77-81.
- Hamerle, A., Liebig, T., & Rösch, D. (2003a). Credit Risk Factor Modeling and the Basel II IRB Approach, Discussion Paper Series 2: Banking and Financial Supervision, No 02/2003, *Deutsche Bundesbank*, available under http://www.bundesbank.de/bank/download/dkp/200302dkp_b.pdf.
- Koyluoglu, H.U., & Hickman, A. (1998b). Reconcilable Differences, *Risk*, 11, October, 56-62.
- Koyluoglu, H.U., & Hickman, A. (1998a). A Generalized Framework for Credit Risk Portfolio Models, Working Paper, *Oliver, Wyman & Co*, available under www.defaultrisk.com.

- Lando, D., & Skodeberg, T. (2002). Analyzing Rating Transitions and Rating Drift with Continuous Observations, *Journal of Banking and Finance*, 26, 423-444.
- Löffler, G. (2004). An Anatomy of Rating Through the Cycle, *Journal of Banking and Finance*, 28, 695-720.
- Rosenblatt, M. (1952). Remarks on a Multivariate Transformation, *The Annals of Mathematical Statistics*, 23, 470-472.
- Rösch, D. (2003). Correlations and Business Cycles of Credit Risk: Evidence from Bankruptcies in Germany, *Financial Markets and Portfolio Management*, 17, 309-331.
- Standard & Poor's (2002). *Corporate Ratings Criteria*, New York, available under www.standardandpoors.com/ratings.
- Standard & Poor's (2001). *Ratings Performance 2000 – Default, Transition, Recovery, and Spreads*. New York, available under www.standardandpoors.com/ratings.
- Treacy, W.F., & Carey, M. (2000). Credit Risk Rating Systems at Large US Banks, *Journal of Banking and Finance*, 24, 167-201.
- Vasicek, O.A. (1991). Limiting Loan Loss Probability Distribution, *KMV Corporation*, San Francisco, available under www.kmv.com; partly reprinted in: Vasicek, O.A. (2002). Loan Portfolio Value, *Risk*, 15, December, 160-162.
- Vasicek, O.A. (1987). Probability of Loss on Loan Portfolio, *KMV Corporation*, San Francisco, available under www.kmv.com; partly reprinted in: Vasicek, O.A. (2002). Loan Portfolio Value, *Risk*, 15, December, 160-162.

Tables and Exhibits

Table 1:

Parameter estimates of model (2) without covariates for grades BB, B, and CCC;
p-values are in parentheses;

_1 means one year time lag, _2 means two years time lag, _chg means one year change of variable

Grade	b	β_0
BB	0.229 (0.003)	-2.290 (<0.001)
B	0.210 (<0.001)	-1.628 (<0.001)
CCC	0.256 (0.004)	-0.809 (<0.001)

Table 2:

Parameter estimates of model (3) with covariates for grades BB, B, and CCC;
p-values are in parentheses;

_1 means one year time lag, _2 means two years time lag, _chg means one year change of variable

FEDR is the Federal Fund Rate, DR_B is the realized default rate of grade B, DSER is the percentage change of real Services, DIND is the percentage change of industry production, RAT4 is the long term bond yield, UNEM is unemployment in percent

Grade	b	β_0	β_1	β_2	β_3	β_4
BB	0.085 (0.333)	-2.791 (<0.001)	0.676 (<0.001) FEDR_1			
B	0.093 (0.010)	-2.104 (<0.001)	3.137 (0.034) DR_B_1	1.453 (<0.001) DSER_2	-0.387 (0.009) DIND_1	7.449 (0.022) RAT4chg_1
CCC	0.048 (0.815)	0.294 (0.264)	-1.798 (<0.001) UNEM_1			

Table 3:

“Naïve” Forecasts of default probabilities (last year’s default rate), issuers and defaults for years 1996-2000 for grades BB, B and CCC; asset correlations are calculated according to Basel II

Grade BB				
Year	Forecast of default probability	Asset Correlation	Issuers	Defaults
1996	0.007	0.205	471	3
1997	0.006	0.207	551	1
1998	0.002	0.230	662	5
1999	0.008	0.202	793	8
2000	0.010	0.193	887	10

Grade B				
Year	Forecast of default probability	Asset Correlation	Issuers	Defaults
1996	0.042	0.135	438	11
1997	0.025	0.154	476	15
1998	0.032	0.145	700	32
1999	0.046	0.132	899	63
2000	0.070	0.124	961	69

Grade CCC				
Year	Forecast of default probability	Asset Correlation	Issuers	Defaults
1996	0.276	0.120	28	1
1997	0.036	0.140	27	3
1998	0.111	0.121	32	11
1999	0.344	0.120	73	22
2000	0.301	0.120	86	25

Table 4:

Forecasts of default probabilities under Through the Cycle Rating, issuers and defaults for years 1996-2000 for grades BB, B and CCC; asset correlations are calculated according to Basel II

Grade BB				
Year	Forecast of default probability	Asset Correlation	Issuers	Defaults
1996	0.012	0.186	471	3
1997	0.011	0.188	551	1
1998	0.010	0.192	662	5
1999	0.010	0.193	793	8
2000	0.010	0.193	887	10

Grade B				
Year	Forecast of default probability	Asset Correlation	Issuers	Defaults
1996	0.053	0.129	438	11
1997	0.050	0.130	476	15
1998	0.048	0.131	700	32
1999	0.048	0.131	899	63
2000	0.051	0.129	961	69

Grade CCC				
Year	Forecast of default probability	Asset Correlation	Issuers	Defaults
1996	0.209	0.120	28	1
1997	0.200	0.120	27	3
1998	0.196	0.120	32	11
1999	0.204	0.120	73	22
2000	0.214	0.120	86	25

Table 5:

Forecasts of default probabilities under Point in Time Rating, issuers and defaults for years 1996-2000 for grades BB, B and CCC; asset correlations are estimated from the model;

Grade BB				
Year	Forecast of default probability	Asset Correlation	Issuers	Defaults
1996	0.008	0.012	471	3
1997	0.007	0.009	551	1
1998	0.006	0.014	662	5
1999	0.006	0.010	793	8
2000	0.006	0.007	887	10

Grade B				
Year	Forecast of default probability	Asset Correlation	Issuers	Defaults
1996	0.038	0.002	438	11
1997	0.026	0.002	476	15
1998	0.049	0.002	700	32
1999	0.049	0.011	899	63
2000	0.098	0.008	961	69

Grade CCC				
Year	Forecast of default probability	Asset Correlation	Issuers	Defaults
1996	0.253	0.004	28	1
1997	0.271	0.005	27	3
1998	0.278	0.021	32	11
1999	0.310	0.014	73	22
2000	0.338	0.006	86	25

Table 6:

LR statistics of the test for adequacy of the forecasted default distributions for “naïve” forecast (last year’s default rate), Through the Cycle and Point in Time Rating for grades BB, B and CCC, p-values are in parentheses

Grade	LR „Naive“	LR Through the Cycle	LR Point in Time
BB	5.94 (0.05)	5.48 (0.06)	1.29 (0.52)
B	6.42 (0.04)	3.24 (0.20)	0.30 (0.86)
CCC	2.30 (0.32)	0.29 (0.87)	4.22 (0.12)

Table 7: Simulated Power of the LR-Test, 10,000 simulations each, $T=5$, default probability=0.01 each, numbers in %

True Rho	Rho under H0	N=1,000			N=10,000		
		alpha=10%	alpha=5%	alpha=1%	alpha=10%	alpha=5%	alpha=1%
0.02	0.01	17.96	10.75	3.22	29.10	19.66	8.19
	0.02 (H0 true)	16.87	9.66	2.68	16.00	9.24	2.51
	0.2	91.18	81.44	48.82	98.95	95.04	68.19
0.05	0.01	39.91	29.99	15.61	74.23	66.56	52.31
	0.05 (H0 true)	15.15	8.59	2.45	15.59	9.16	2.38
	0.2	67.19	50.41	21.27	73.98	57.61	24.90
0.1	0.01	70.36	61.81	44.45	94.16	92.05	86.51
	0.1 (H0 true)	16.24	9.60	2.49	14.83	8.31	2.19
	0.2	38.49	25.00	8.04	37.37	23.96	7.79
0.2	0.01	91.80	87.78	77.81	99.35	99.03	98.21
	0.2 (H0 true)	17.96	10.82	3.03	14.99	8.59	2.38

Table 8: Simulated Power of the LR-Test, 10,000 simulations each, $T=5$, default probability=0.05 each, numbers in %

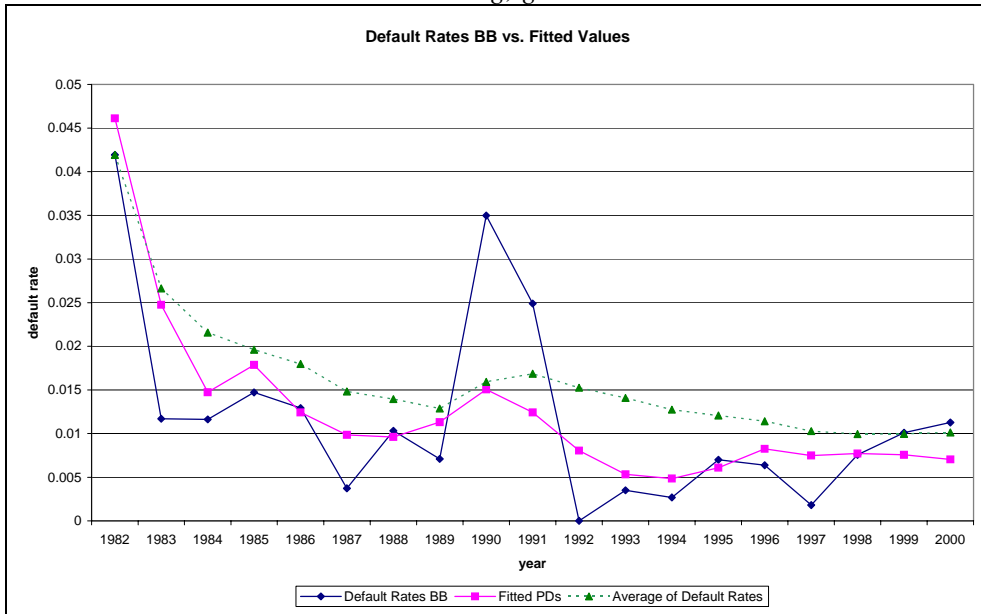
True Rho	Rho under H0	N=1,000			N=10,000		
		alpha=10%	alpha=5%	alpha=1%	alpha=10%	alpha=5%	alpha=1%
0.02	0.002	63.96	55.17	39.24	90.29	87.37	80.20
	0.02 (H0 true)	16.41	9.25	2.38	15.61	9.02	2.36
	0.13	79.99	63.47	29.04	86.33	71.98	35.35
0.05	0.002	91.42	88.03	80.99	98.66	98.06	96.86
	0.05 (H0 true)	15.05	8.66	2.42	15.79	9.08	2.50
	0.13	40.00	26.16	8.37	41.69	27.51	9.12
0.1	0.002	98.15	97.32	95.05	99.81	99.68	99.50
	0.1 (H0 true)	15.44	8.99	2.45	15.20	8.80	2.19
	0.13	18.03	10.52	3.03	18.26	10.48	2.97
0.13	0.002	99.05	98.68	97.58	99.93	99.89	99.79
	0.13 (H0 true)	14.89	8.32	2.27	15.90	9.63	2.61

Table 9:

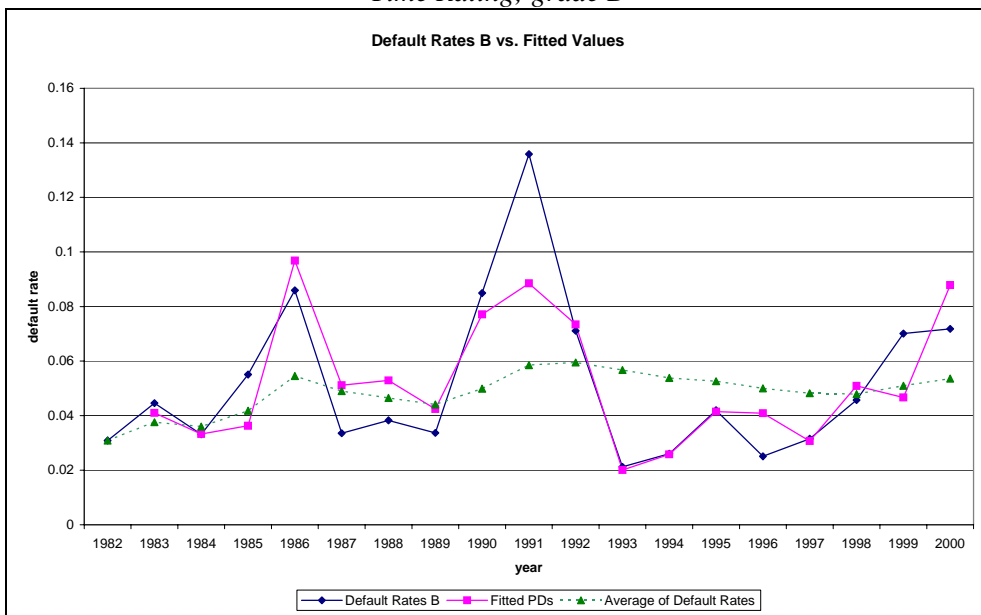
Value-at-Risk quantiles of forecasted default distributions for different rating grade; N is the number of borrowers in the portfolio, defaults are as percentages of portfolio size; the upper half shows the quantiles of the model which assumes Through the Cycle probability of default under Basel II correlation, the lower half shows the quantiles of the model with Point in Time probability of default and correlation; Value-at-Risk is defined as a quantile of the default distribution; Unexpected Loss is defined as the difference between the Value-at-Risk quantile and the expected loss; VaR0 (UL0) is the Value-at-Risk (Unexpected Loss) with Through the Cycle probability of default and Basel II asset correlation, VaR1(UL1) is the Value-at-Risk (Unexpected Loss) with Point in Time default probability and correlation; numbers are in %

Portfolio size	Grade BB N=887			Grade B N=961			Grade CCC N=86		
	99	99.5	99.9	99	99.5	99.9	99	99.5	99.9
Confidence Level (%)									
VaR0	7.33	9.13	13.98	19.46	22.16	28.51	50.00	53.49	61.63
EL0		1.00			5.09			19.23	
UL0	6.33	8.13	12.98	14.37	17.07	23.42	30.77	34.26	42.40
VaR1	1.47	1.58	1.80	14.57	15.09	16.34	47.67	48.84	52.33
EL1		0.64			9.79			33.79	
UL1	0.83	0.94	1.16	4.78	5.30	6.55	13.88	15.05	18.54

*Exhibit 1:
Default rates and probability of default (PD) estimates due to Through the Cycle and Point in
Time Rating; grade BB*



*Exhibit 2:
Default rates and probability of default (PD) estimates due to Through the Cycle and Point in
Time Rating; grade B*



*Exhibit 3:
Default rates and probability of default (PD) estimates due to Through the Cycle and Point in
Time Rating; grade CCC*

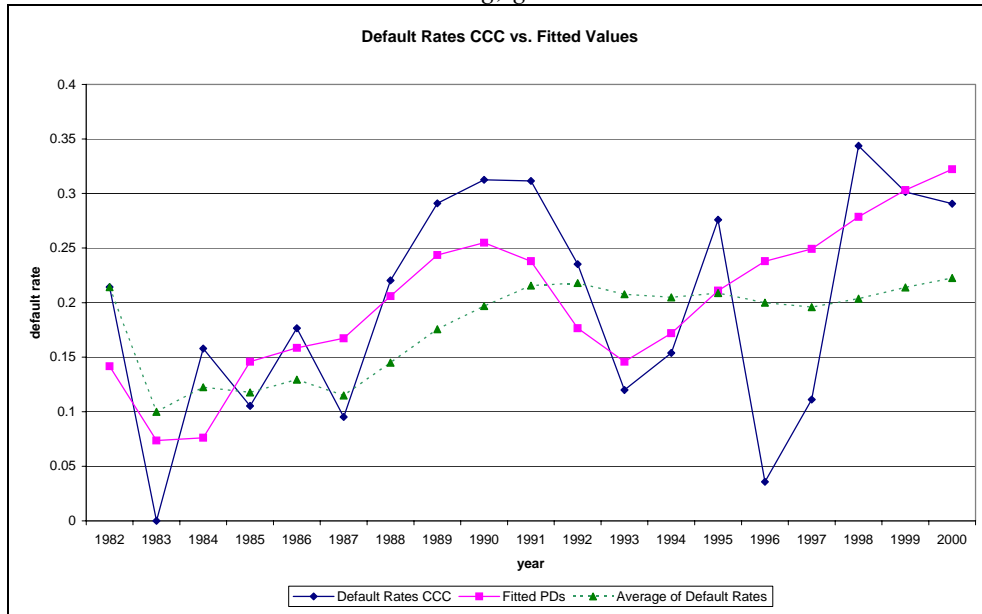


Exhibit 4:
Forecasted Default Distributions for year 2000; Through the Cycle versus Point in Time Rating; grade BB

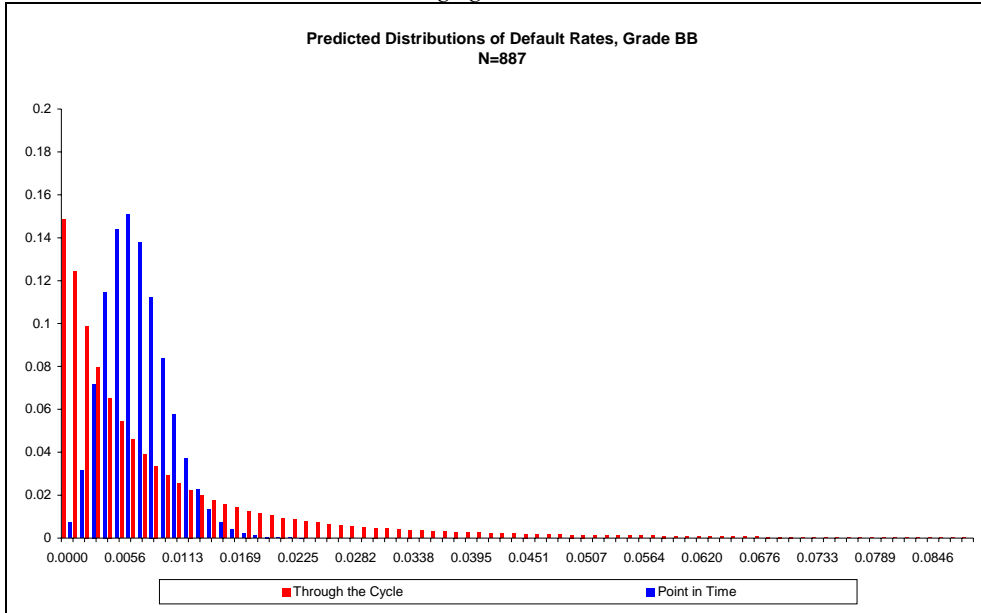
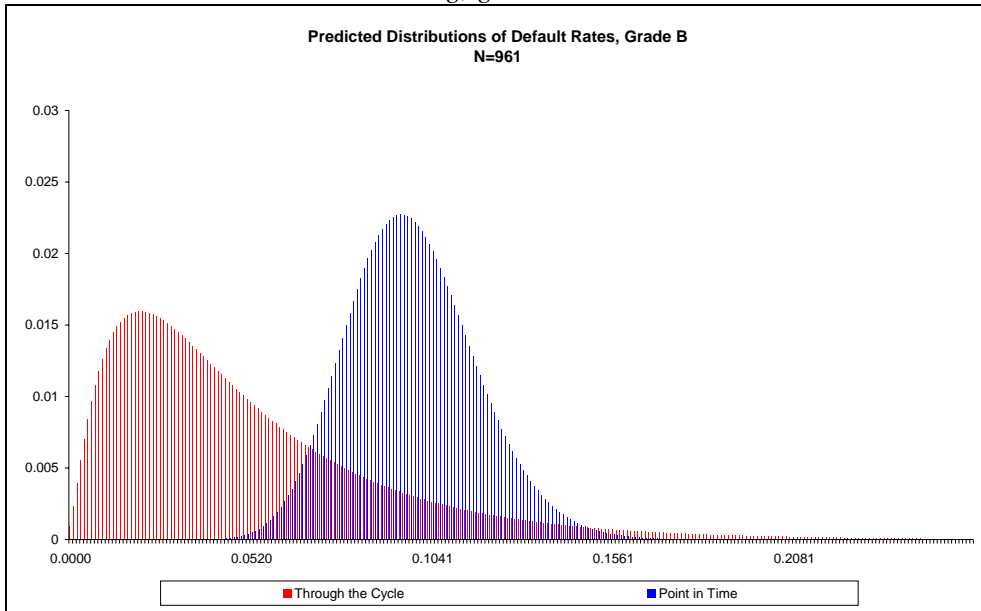


Exhibit 5:
Forecasted Default Distributions for year 2000; Through the Cycle versus Point in Time Rating; grade B



*Exhibit 6:
Forecasted Default Distributions for year 2000; Through the Cycle versus Point in Time Rating; grade CCC*

