IFRS9 PD modeling Presentation MATFYZ 03.01.2018 Konstantin Belyaev

IFRS9 International Financial Reporting Standard

- In July 2014, the International Accounting Standard Board (IASB) issued the final version of **IFRS 9** Financial Instruments, bringing together the classification and measurement, **impairment** and hedge accounting phases of the IASB's project to replace IAS 39 and all previous versions of **IFRS 9**.
- The standard brings together three phases:
  - Phase 1: Classification and measurement
  - **Phase 2: Impairment methodology**
  - Phase 3: Hedge accounting

Live from 1 January 2018

IFRS9 principles

Basel II / IAS39	IFRS9
Reflects economic cycle - **stability**	Aligned with accounting view- **volatility**
TTC and **DT** view	**PIT** estimates, macro-economic environment, including forecasts
Estimates EL for **1-year**	**Multi-year** aspect
IAS39 – **incurred losses**	**Expected** credit loss
Conservative	**Best estimate**

Multi-year PD modeling

- For the sake of clarity all illustrations of approaches presented here, make the assumption that a rating scale is available (Numerical ratings 1-12, or Letters AAA, BBB, C, D)

Approach	Description
A Markov chain based approaches	**A1: Homogeneous Discrete-time Markov Chain Method** ▪ Estimate cumulative PD profiles by means of a migration matrix. ▪ The cumulative migration probabilities of the migration matrix are estimated by means of the cohort method and therefore only for discrete time slices.
	A2: Homogeneous Continuous-time Markov Chain Method ▪ Estimate cumulative PD profiles by means of a generator (for multi-year migration matrices). ▪ The cumulative migration probabilities of the migration matrix are estimated by means of the cohort method. The discrete-time matrices are transformed to generators.
	A3: Non-Homogeneous Continuous Markov Chain Method ▪ Estimate cumulative PD profile by means of different migration matrices for different time periods. ▪ The method is also based on generators but the time component will be modelled so that the default rates are approximated.
B Survival analysis based approaches	**B1: Weibull Survival Probability Method** ▪ Estimate cumulative PD profiles based on internal default histories or default histories available from external providers (e.g., S&P's). ▪ Estimates the Weibull fitting parameters k and λ by means of a maximum likelihood estimation (MLE).
	B2: Weibull Fitting on Historical Default Rates ▪ Estimate cumulative PD profiles based on internal default histories or default histories available from external providers (e.g., S&P's). ▪ Estimates the Weibull fitting parameters k and λ by means of a linear regression on the double logarithm of the survival function

Next year	This year	1Y-migration matrix
Next year's debtors' distribution stored in vector: $\vec{r}_{next\_year}$	This year's debtors' distribution stored in vector: $\vec{r}_{this\_year}$	Observed transition rates from rating i to rating j within a year are stored as entries $m_{i,j}$ of the transition matrix $\boldsymbol{M_1}$. In particular, the default probability for a debtor in class i over the coming year is given by $m_{i,n}$.

1Y-migration formalism

$$\vec{r}_{next\ year} = \vec{r}_{this\_year} \cdot \begin{pmatrix} m_{1,1} & m_{1,2} & \dots & m_{1,n} \\ \dots & \dots & \dots & \dots \\ m_{n-1,1} & m_{n-1,2} & \dots & m_{n-1,n} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Markov property assumption

- The future rating transitions depend only on the current rating but not on any previous ratings.

Homogeneity assumption

- The migration probabilities $m_{i,j}$, do NOT depend on the specific point in time, i.e. the transition rates $m_{i,j}$ do not change with time t.

t-th year cumulative PDs

- Under these assumptions, a t-year transition matrix can be determined straightforwardly as the t^{th} power of the one-year transition matrix:

$$M_t \coloneqq M_1^t$$

- The last column of the transition matrix, $(M_t)_n$, contains the t-year cumulative PDs (CPDs).

[1] Number of debtors with rating i is stored in the i-th element of a vector $\vec{r} = (r_1, \dots, r_i, \dots, r_n)$

Cohort method

Method

- First order Markov-Chain in discrete time
- Migration matrix

$$M = (m_{ij}) = \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1n} \\ \dots & \dots & \dots & \dots \\ m_{n-1,1} & m_{n-1,2} & \dots & m_{n-1,n} \\ 0 & 0 & \dots & 1 \end{bmatrix}; \quad M(t) = M^t$$

- Estimator for $m_{ij} = P(x_{t+1} \text{ is } j \mid x_t = i)$

$$\hat{m}_{ij} = \frac{N_{ij}(t)}{N_i(t-1)} = \frac{\text{Amount of migrations from i to j in period t-1 until t}}{\text{Amount of borrowers in i at time t-1}}$$

Proper-ties

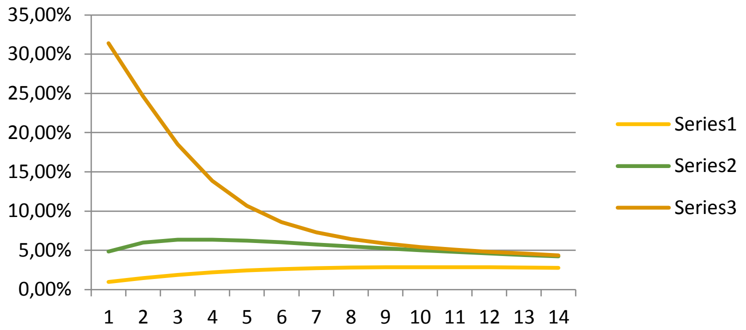
- Time homogeneity assumption
- Does not take into account migrations occurring within one period
- Not-observed migrations receive probabilities of zero, e.g., extreme migration AAA → D
- If applicable, no economic plausible results, e.g., no monotonous probabilities of default across rating classes

Duration method

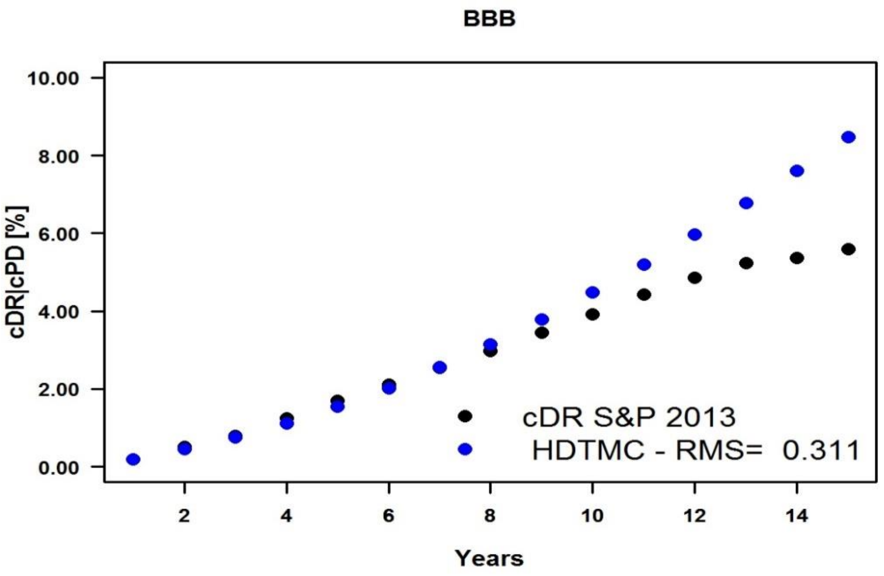
Pros:
- Use of all available information (incl. intra-annual rating changes)
- Delivers non-zero default probabilities even when no actual defaults were observed

Cons:
- Typically overestimates (underestimates) default probabilities in rating classes C/CCC (all other rating classes)
- High data requirements and lower transparency

Desirable properties	Description
Monotonicity	**Dominance of cumulative transition probabilities over rating classes** ■ For any rating class k it should hold that moving to this or a better class is more likely, the higher (resp. the worse) the original rating class: $$\sum_{j \le k} m_{i,j} \ge \sum_{j \le k} m_{i+1,j}, \qquad \forall k$$
Unimodality	**Unimodality of transition probabilities over rating classes** ■ For any rating class k, it should hold that the transition probabilities are distributed unimodally, with probabilities that are falling monotonically when moving away from the diagonal. Otherwise, sudden jumps in creditworthiness would be possible, while from an economic-financial point of view a smooth behavior is typically preferred.



Dominance of the forward PDs	The forward PD term structure for a given rating class should dominate those of worse rating classes ■ Conditional PDs for good rating classes should never be higher than those for worse rating classes for all future periods.
Fitting performance	Cumulative PDs should provide reasonable fit to observed cumulative default rates ■ The cumulative probabilities of default, represented in the last column of the cumulative transition matrices, should fit the observed multi-year default rates reasonably well.



- **Weaknesses:**
  - By means of the cohort method, cumulative transition probabilities are estimated only for discrete time slices of a fixed length (i.e. 1y, 2y, etc.). In reality, inter-annual CPDs have to be estimated for most of the transactions.
  - Long-run cDR (e.g., >9Y) are poorly fitted by the estimated CPD.
- **Possible solutions:**
  - Transform the discrete-time matrices to continuous-time matrices.
  - Test alternative estimation approaches.

Cumulative DR (cDR) and estimated cumulative PD (CPD) values for rating classes AAA, BBB and B for years 1 to 15 after initial rating

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
AAA - HDTMC	0.01%	0.04%	0.07%	0.12%	0.17%	0.23%	0.30%	0.37%	0.44%	0.53%	0.62%	0.72%	0.83%	0.95%	1.07%
AAA - Emp	0.00%	0.03%	0.13%	0.24%	0.35%	0.47%	0.53%	0.62%	0.68%	0.74%	0.77%	0.81%	0.84%	0.91%	0.99%
BBB - HDTMC	0.2%	0.5%	0.8%	1.1%	1.5%	2.0%	2.6%	3.1%	3.8%	4.5%	5.2%	6.0%	6.8%	7.6%	8.5%
BBB - Emp	0.2%	0.5%	0.8%	1.2%	1.7%	2.1%	2.6%	3.0%	3.4%	3.9%	4.4%	4.9%	5.2%	5.4%	5.6%
B - HDTMC	4.7%	10.4%	16.3%	22.0%	27.4%	32.3%	36.8%	40.9%	44.5%	47.8%	50.8%	53.5%	55.9%	58.1%	60.2%
B - Emp	4.7%	10.6%	15.2%	18.5%	21.0%	23.3%	24.8%	25.8%	26.8%	27.7%	28.5%	29.3%	30.0%	30.6%	31.4%

Idea: Transform the discrete-time matrices to continuous-time matrices by using the matrix exponential function that gives a continuous generalization for powers of a matrix.

- M: 1-year migration matrix
- G: generator matrix
- exp(·) denotes the matrix exponential and $t \geq 0$

$$M = \exp(G) \quad \Rightarrow CPD_i(t) = (\exp(G \cdot t))_{i,n}$$

Example algorithm:

- Compute the logarithm of M:

$$\hat{G} = \ln(M)$$

- Check generator matrix requirements:
  - Diagonal entries not positive: $\widehat{g_{i,i}} \leq 0,\ i = 1, \dots, n$
  - Off –diagonal entries not negative: $\widehat{g_{i,j}} \geq 0,\ i, j = 1, \dots, n \text{ and } i \neq j$
  - Sum of entries of each row is 0: $\sum_{j=1}^{n} \widehat{g_{i,j}} = 0,\ i = 1, \dots, n$

- Weaknesses:

In many practical cases, the 1-year migration matrix does not have a regular generator matrix.

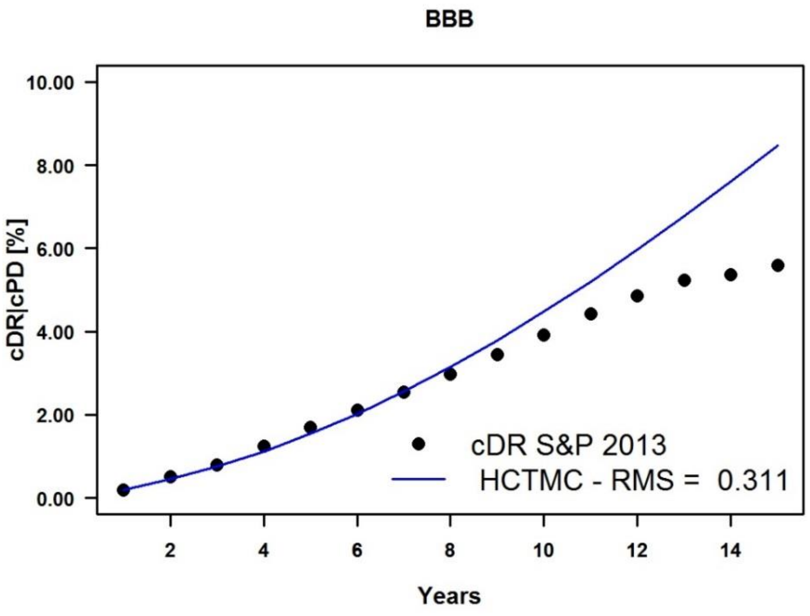
- Possible solutions:

Application of a so-called regularization algorithm (cf. for example Israel et al. [2001] and Kreinin and Sidelnikova [2001]).

- Regularization algorithm technique examples:
  - Replace all negative non-diagonal entries of G by zero:

$$\widehat{g_{i,j}} = \begin{cases} 0 \text{ if } i \neq j \text{ and } g_{i,j} < 0 \\ g_{i,j}, \text{otherwise} \end{cases}, \qquad i,j = 1,\dots,n$$

  - Adjust elements to ensure that each row sums to zero
    - Diagonal adjustment: $\widehat{g_{i,i}} = -\sum_{j=1, j\neq i}^{n} \widehat{g_{i,j}}$, $i = 1,\dots,n$
    - Weighted adjustment: $\widehat{g_{i,j}} = \widehat{g_{i,j}} - \left|\widehat{g_{i,j}}\right| * \frac{\sum_{i=1}^{n} \widehat{g_{i,j}}}{\sum_{i=1}^{n} \left|\widehat{g_{i,j}}\right|}$, $i,j = 1,\dots,n$



- Weaknesses:

  Continuous CPD forecasts show systematic overestimation for long time horizons (e.g., t>9y).

- Possible solutions:

  Use a non-homogeneous continuous time migration matrix method, i.e. use different migration matrices for $t_1 \rightarrow t_1 + \Delta t$ and $t_2 \rightarrow t_2 + \Delta t$.

Idea: Allow for time-dependent migrations by means of a time-dependent modification of the generator matrix [1]

$$Q_t \equiv \begin{pmatrix} \varphi_{\alpha_1,\beta_1}(t) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \varphi_{\alpha_n,\beta_n}(t) \end{pmatrix}_{(n,n)} \times G$$

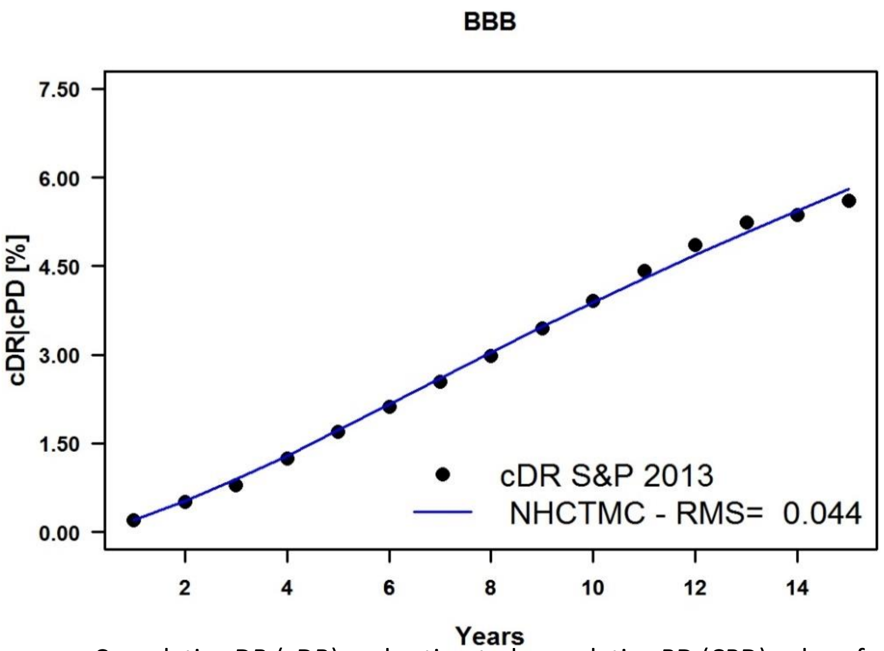
- Where:
  - Q_t: the modified $(n \text{ x } n)$ generator matrix,
  - G: the $(n \times n)$ Homogeneous Continuous Time Migration Matrix Method generator
  - $\varphi_{\alpha_i,\beta_i}(t) \equiv \frac{(1-e^{-\alpha_i t})}{1-e^{-\alpha_i}} \cdot t^{\beta_i - 1}$: time and rating class dependent modification functions;
    α_i and β_i are used to fit the empirical cDRs

- Estimation of fitting parameters α_i and β_i (for T years of cDR data):

$$(\alpha_i, \beta_i) = \underset{(\alpha_i, \beta_i)}{\operatorname{argmin}} \left(\sqrt{\frac{1}{T} \sum_{t=1}^{T} \left(cDR_{i,t} - \widehat{CPD}_{i,t} \right)^2} \right)$$, where $\widehat{CPD}_{i,t}$ is the estimated cumulative PD

- $\varphi_{\alpha_n,\beta_n}(t)$ can be interpreted as decelerated or accelerated, so called statistical time

1) Bluhm, C. and Overbeck, L., (2007): “To be Markovian or not to be”, Risk: managing risk in the world’s financial markets, Vol. 20 (11), pp. 98-103.



Cumulative DR (cDR) and estimated cumulative PD (CPD) values for rating classes AAA, BBB and B for years 1 to 15 after initial rating

- **Strengths:**

  Non-homogeneous continuous time Markov chain method results in a close fit of the empirical data.

- **Weaknesses:**

  No theoretical foundation of the time adjustment.

  The whole migration matrix as well as two parameters α_i and β_i per rating class are required. For N different rating classes, the total number of parameters therefore is:

$$2 \cdot n + n \cdot (n-1) = n^2 + n$$

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
AAA - NHCTMC	0.01%	0.05%	0.11%	0.18%	0.25%	0.33%	0.41%	0.49%	0.58%	0.66%	0.75%	0.84%	0.93%	1.02%	1.11%
AAA - Emp	0.00%	0.03%	0.13%	0.24%	0.35%	0.47%	0.53%	0.62%	0.68%	0.74%	0.77%	0.81%	0.84%	0.91%	0.99%
BBB - NHCTMC	0.2%	0.5%	0.9%	1.3%	1.7%	2.2%	2.6%	3.0%	3.5%	3.9%	4.3%	4.7%	5.1%	5.4%	5.8%
BBB - Emp	0.2%	0.5%	0.8%	1.2%	1.7%	2.1%	2.6%	3.0%	3.4%	3.9%	4.4%	4.9%	5.2%	5.4%	5.6%
B - NHCTMC	4.7%	10.1%	14.5%	17.9%	20.5%	22.6%	24.3%	25.7%	26.9%	27.9%	28.8%	29.6%	30.3%	31.0%	31.6%
B - Emp	4.7%	10.6%	15.2%	18.5%	21.0%	23.3%	24.8%	25.8%	26.8%	27.7%	28.5%	29.3%	30.0%	30.6%	31.4%

Idea: The question whether and when a client defaults could be seen as a survival process.

In survival theory, a widely used[3] survival function is:

$$S(t) \coloneqq 1 - F(t;\kappa,\lambda),$$

where $F(t;\kappa,\lambda)$ denotes the 2-parameter Weibull distribution function.

$$F(t;\kappa,\lambda) = \begin{cases} 1 - e^{-\left(\frac{t}{\lambda}\right)^{\kappa}}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

- Where:
  - k > 0 controls the overall shape of the density function.
    - k < 1 indicates that the default rate decreases over time
    - k = 1 indicates that the default rate is constant over time
    - k > 1 indicates that the default rate increases over time
  - Typically, k ranges between 0.5 and 8.0
  - The scale parameter λ > 0 controls the survival time; for $t = \lambda$ the CPD is $1 - e^{-1}$ ≈ 63%

Either increasing or decreasing

- The probability density function of a Weibull random variable is given by:

$$f(\kappa,\lambda,t) = \begin{cases} \frac{\kappa}{\lambda} \cdot \left(\frac{t}{\lambda}\right)^{\kappa-1} \cdot e^{-\left(\frac{t}{\lambda}\right)^{\kappa}}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

Principle assumption

- The CPD is represented by a Weibull distribution:

$$CPD(t) = 1 - e^{-\left(\frac{t}{\lambda}\right)^{k}} = F(t; \kappa, \lambda)$$

MLE - Method

- Weibull fitting parameters k and λ are determined by maximization of the Log-Likelihood

Truncation

- Two types of default data in empirical credit default database:
  - Uncensored data: credits that defaulted during the observation period
  - Right-censored data: credits that fully survived the observation period

Weighting

- $f(k, \lambda, t_i)$ for an uncensored observation, where t_i represents the survival time (e.g., default date minus rating attribution date)
- $1 - F(k, \lambda, t_i)$ for a right-censored observation, where t_i represents the truncated rating duration (e.g., today’s date minus rating attribution date)

Determine parameters

- With δ_i as „Truncation-Indicator” (e.g., 0 if uncensored and 1 if right censored) the Log-Likelihood is:

$$LL = \sum_{i=1}^{N} \left(\delta_i \ln\left(1 - F(\kappa, \lambda, t_i)\right) + (1 - \delta_i) \ln\left(f(\kappa, \lambda, t_i)\right) \right)$$

As the Weibull Survival Probability method is based on MLE, it is a well-defined approach with advantageous properties.

Principle assumption

- The CPD is assumed to be represented by means of a Weibull distribution function:

$$\mathrm{CPD}(t) = 1 - e^{-\left(\frac{t}{\lambda}\right)^{k}} = F(t; \kappa, \lambda)$$

Transform to a linear model

- Survival function: $S(t) = 1 - \mathrm{CPD}(t) = e^{-\left(\frac{t}{\lambda}\right)^{k}}$

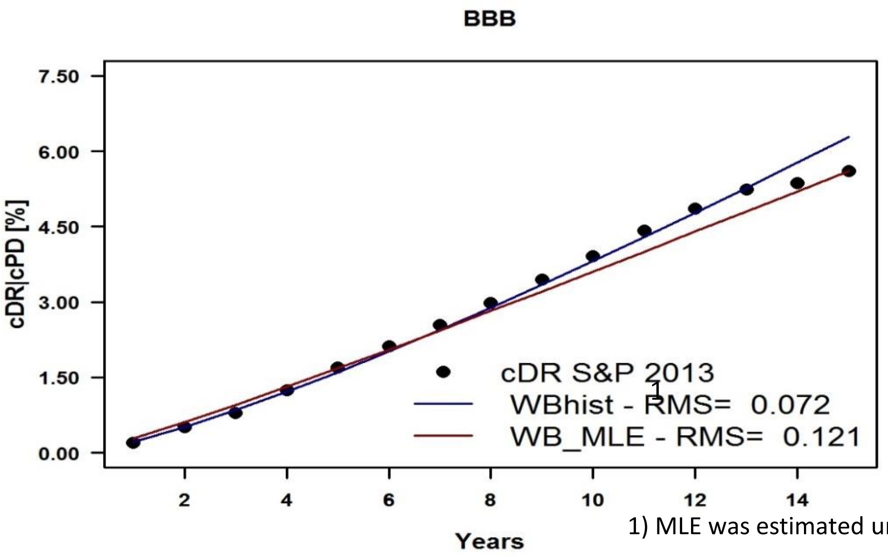
$$\ln\big(S(t)\big) = -\left(\frac{t}{\lambda}\right)^{k}$$

$$\Leftrightarrow \ln\!\big(-\ln\big(S(t)\big)\big) = b * \ln(t) + a$$

where: $\kappa = b$ and $\lambda = e^{-a/\kappa}$ are estimated at rating class level

Determine parameters by linear regression

- The linear relationship between $\ln(t)$ and $\ln\Big(-\ln\big(S(t)\big)\Big)$ allows to obtain the Weibull parameters κ and λ by means of a linear regression.



- **Strengths:**

  Using a common distribution to model survival probabilities is a sound theoretical foundation

  Low implementation effort

- **Weaknesses:**

The model performs much better than HCTMC but compared to NHCTMC, the results show…

… poor extrapolation

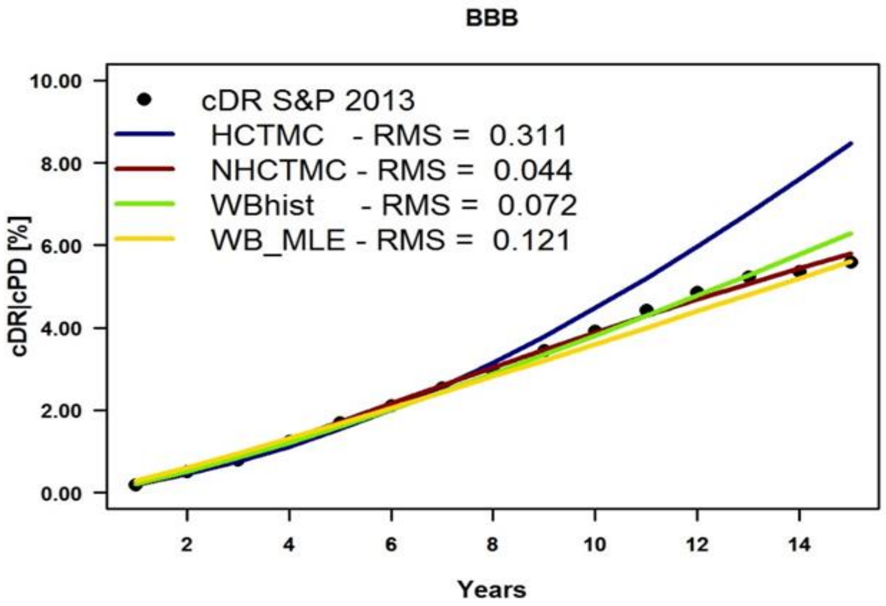
… only moderate goodness of fitting

Survival rates might not be Weibull distributed

1) MLE was estimated under the assumption that survival rates are given by 1-DR

Cumulative DR (cDR) and estimated cumulative PD (CPD) values for rating classes AAA, BBB and B for years 1 to 15 after initial rating

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
AAA - WBhist	0,02%	0,06%	0,12%	0,18%	0,25%	0,33%	0,42%	0,51%	0,60%	0,71%	0,81%	0,93%	1,04%	1,17%	1,29%
AAA - WB_MLE	0,05%	0,10%	0,16%	0,22%	0,29%	0,35%	0,42%	0,49%	0,56%	0,63%	0,70%	0,77%	0,84%	0,92%	0,99%
AAA - Emp	0,00%	0,03%	0,13%	0,24%	0,35%	0,47%	0,53%	0,62%	0,68%	0,74%	0,77%	0,81%	0,84%	0,91%	0,99%
BBB - WBhist	0,2%	0,5%	0,8%	1,2%	1,6%	2,0%	2,4%	2,9%	3,3%	3,8%	4,3%	4,8%	5,3%	5,8%	6,3%
BBB - WB_MLE	0,3%	0,6%	1,0%	1,3%	1,7%	2,1%	2,4%	2,8%	3,2%	3,6%	4,0%	4,4%	4,8%	5,2%	5,6%
BBB - Emp	0,2%	0,5%	0,8%	1,2%	1,7%	2,1%	2,6%	3,0%	3,4%	3,9%	4,4%	4,9%	5,2%	5,4%	5,6%
B - WBhist	6,5%	10,2%	13,3%	16,0%	18,4%	20,6%	22,6%	24,5%	26,3%	28,0%	29,6%	31,1%	32,5%	33,9%	35,2%
B - WB_MLE	6,9%	10,4%	13,2%	15,6%	17,7%	19,5%	21,3%	22,9%	24,4%	25,8%	27,1%	28,4%	29,6%	30,7%	31,8%
B - Emp	4,7%	10,6%	15,2%	18,5%	21,0%	23,3%	24,8%	25,8%	26,8%	27,7%	28,5%	29,3%	30,0%	30,6%	31,4%

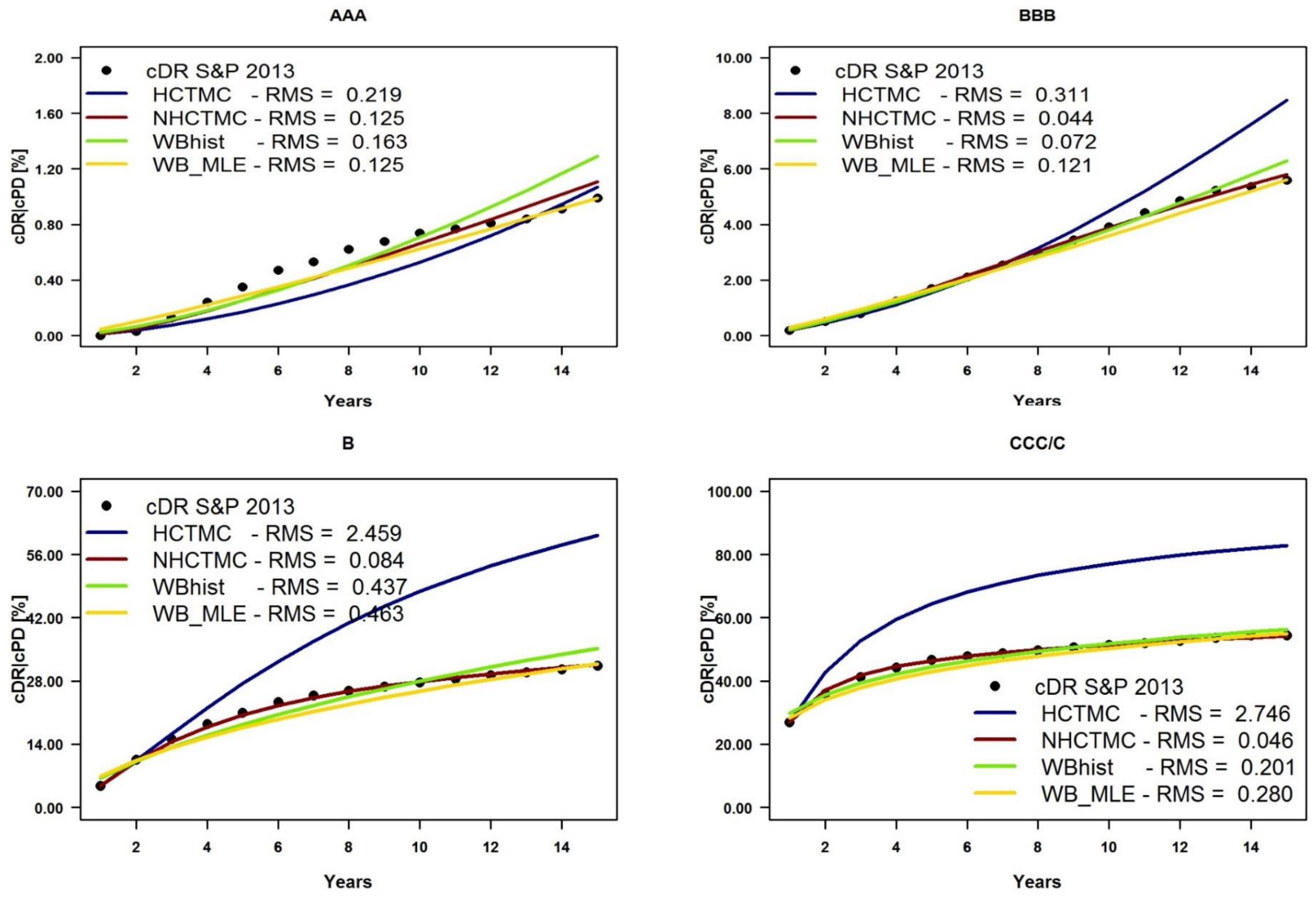


Discussion

- All fitting methods remedy the systematic overestimation of the HCTMC
- NHCTMC performs best (within the sample) but it also has the most parameters
- Weibull methods show poorer fit in the short run than in the long run (better visible on the next slide)

Further discussion

- All methods are fitting methods – for out-of-sample or extrapolation application one has to consider
  - Stability of fitting parameters for different data
  - Long term behavior of fitting curves
  - Extrapolation behavior (are there systematic over-/underestimations, e.g., due to changes in the empirical data slope or curvature)
- Incorporation of macro-economic forecasts


AAA
cDR S&P 2013
HCTMC - RMS = 0.219
NHCTMC - RMS = 0.125
WBhist - RMS = 0.163
WB_MLE - RMS = 0.125
cDR|cPD [%]
Years
BBB
cDR S&P 2013
HCTMC - RMS = 0.311
NHCTMC - RMS = 0.044
WBhist - RMS = 0.072
WB_MLE - RMS = 0.121
cDR|cPD [%]
Years
B
cDR S&P 2013
HCTMC - RMS = 2.459
NHCTMC - RMS = 0.084
WBhist - RMS = 0.437
WB_MLE - RMS = 0.463
cDR|cPD [%]
Years
CCC/C
cDR S&P 2013
HCTMC - RMS = 2.746
NHCTMC - RMS = 0.046
WBhist - RMS = 0.201
WB_MLE - RMS = 0.280
cDR|cPD [%]
Years