A Latent Variable Approach to Validate Credit Rating Systems using R

Chicago, April 24, 2009

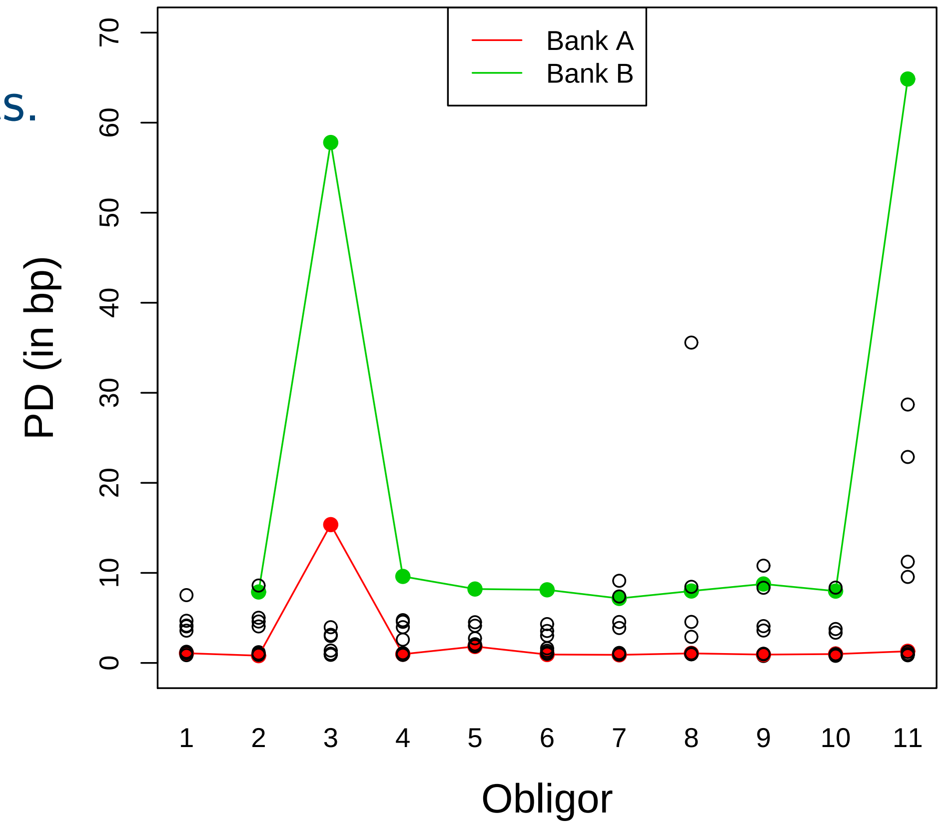
Bettina Grün[a], Paul Hofmarcher[a], Kurt Hornik[a], Christoph Leitner[a], Stefan Pichler[a]

[a] WU Wien

Motivation I

Multi-rater panel provides information about rating heterogeneity

- ✔ across obligors,
- ✔ across raters,
- ✔ across time points.



Motivation II

- ✔ Rating validation:
  - ✘ How accurate (or precise) are PD estimates?
  - ✘ Measures of rating bias and “error variance”.
- ✔ Consensus PD:
  - ✘ What is the consensus PD estimate when ratings differ across raters?
- ✔ Explaining rating heterogeneity:
  - ✘ Opaqueness of industries (Morgan 2002),
  - ✘ Regional differences (Hornik et al. 2009).

Approach

- ✔ Need for a probabilistic model.
- ✔ Our Approach: PD treated as latent variable
  - ✘ Ratings are noisy observations of PD.
  - ✘ Provides a probabilistic model to directly estimate the rating error of a rater.
  - ✘ Explore patterns of heterogeneity.
  - ✘ Can obtain the “consensus rating” of the obligor.
- ✔ Requirement: Multi rater panel with rating information on the same metric scale.
  - ✘ In practice: “PD ratings” (PD estimates by different raters for all obligors).

General Model I

WIRTSCHAFTS UNIVERSITÄT WIEN

- Merton (1974): Distance to default of firm i

$$DD_i = a_i + b_i V_i$$

where V_i is the log asset value of firm i.

- The probability of default is given as $PD_i = \Phi(-DD_i)$ where

$$\Phi : (-\infty, +\infty) \to [0, 1]$$

is the distribution function of the standard normal distribution.

- Probit of the PD is linearly related to the log asset value.

General Model II

- The "estimate" PD_{ij} as derived by rater j for the true PD of firm i is

$$PD_{ij} = \Phi(-DD_i + \epsilon_{ij})$$

  where ϵ_{ij} is the corresponding error (which depends on j in particular).

- Motivates a multi-rater model where the rater effects ("errors") are additive on the probit (Φ^{-1}, DD_i) scale, i.e.,

$$\Phi^{-1}(PD_{ij}) = -DD_i + \epsilon_{ij} = \Phi^{-1}(PD_i) + \epsilon_{ij}$$

- Similar considerations apply for PD models based on, e.g., logistic regression.

General Model III

- ✔ Let $S_{ij} = \Phi^{-1}(PD_{ij})$ denote the *score* of the observed PD and $S_i = \Phi^{-1}(PD_i)$ the latent true PD score.
- ✔ Write μ_{ij} and σ_{ij} for the mean and standard deviation of ϵ_{ij}, respectively.
- ✔ The latent variable model can be written as

$$S_{ij} = S_i + \mu_{ij} + \sigma_{ij} Z_{ij}$$

where the standardized rating errors $Z_{ij} = (\epsilon_{ij} - \mu_{ij})/\sigma_{ij}$ have mean zero and unit variance.

WIRTSCHAFTS UNIVERSITÄT WIEN

Specific Models I

WIRTSCHAFTS UNIVERSITÄT WIEN

- ✔ Simple model:

$$S_{ij} = S_i + \mu_j + \sigma_j Z_{i,j}$$

  - ✘ Rating errors are independent of obligors and their characteristics (e.g., creditworthiness).
  - ✘ Parameters depend on raters only: $\mu_{ij} = \mu_j, \sigma_{ij} = \sigma_j$.

- ✔ Extended model:

$$S_{ij} = S_i + \mu_{g(i),j} + \sigma_{g(i),j} Z_{ij}$$

  - ✘ Groups of obligors with homogeneous rating error characteristics.
  - ✘ Parameters depend on rater and obligor groups: $\mu_{ij} = \mu_{g(i),j}, \sigma_{ij} = \sigma_{g(i),j}$.

Specific Models II

WIRTSCHAFTS UNIVERSITÄT WIEN

- ✔ Ensuring identifiability:

  Average rating bias within each obligor group is zero.

  $\mathbb{E}(S_{ij}) = \mathbb{E}(S_i) + \mu_{g(i),j}$ with $\sum_j \mu_{g,j} = 0$.

- ✔ Modelling marginal group effects:

  $\mathbb{E}(S_i) = \nu_{g(i)}$

  $\mathrm{var}(S_i) = \tau^2$ or $\mathrm{var}(S_i) = \tau^2_{g(i)}$.

Estimation

- ✔ Mixed-effects model for the observed S_{ij},
  - ✘ latent true PD scores S_i enter as random effects.
- ✔ Given parametric models for the μ_{ij} and σ_{ij} and the distributions of true PD scores and standardized rating errors.
- ✔ Estimation using R:
  - ✘ Marginal maximum likelihood: nlme (ratermodel), or
  - ✘ Bayesian techniques: rjags and coda

Offerings

- ✔ Assessment of rater effects (rater heterogeneity) in terms of their μ/σ^2 bias/variance characteristics.
- ✔ Estimated random effects $\hat{S}_i$ can be interpreted as the consensus rating for obligor i.
- ✔ Residuals: $S_{ij} - \hat{S}_i - \hat{\mu}_{ij}$, the observed scores minus the estimated random and fixed effects,
  - ✘ are analyzed using some covariates (e.g., country, industry, legal form),
  - ✘ are analyzed for single obligors as well as obligor groups of interest.

Data

- ✔ Data from the Austrian Credit Register from September 2007.
- ✔ Multi-rater panel containing:
  - ✘ obligor specific information (country, industry, ...),
  - ✘ rating information (original rating, PD).
- ✔ Include 13 banks which report PDs and have sufficiently many co-ratings (every bank has at least 4 co-ratings with every other bank).
  - ✘ 5432 co-ratings of 13 banks for 2090 obligors.

WIRTSCHAFTS UNIVERSITÄT WIEN

Model Selection

- ✔ Variety of models analyzed using AIC and BIC.
  - ✘ Industry affiliation as grouping variable.
  - ✘ Constant variance τ^2 for PD score.

$$S_{ij} = S_i + \mu_{g(i),j} + \sigma_{g(i),j} Z_{ij}, \qquad S_i \sim N(\nu_{g(i)}, \tau^2),$$

$$\mathbb{E}(S_{ij}) = \mathbb{E}(S_i) + \mu_{g(i),j}, \qquad \sum_j \mu_{g,j} = 0,$$

where

$\mu_{g,j}$ .... rating bias to the mean PD score of bank j for obligors in industry g,

$\sigma_{g,j}$ .... standard deviation of the rating error of bank j for obligors in industry g, and

ν_g ...... mean PD score in industry g.

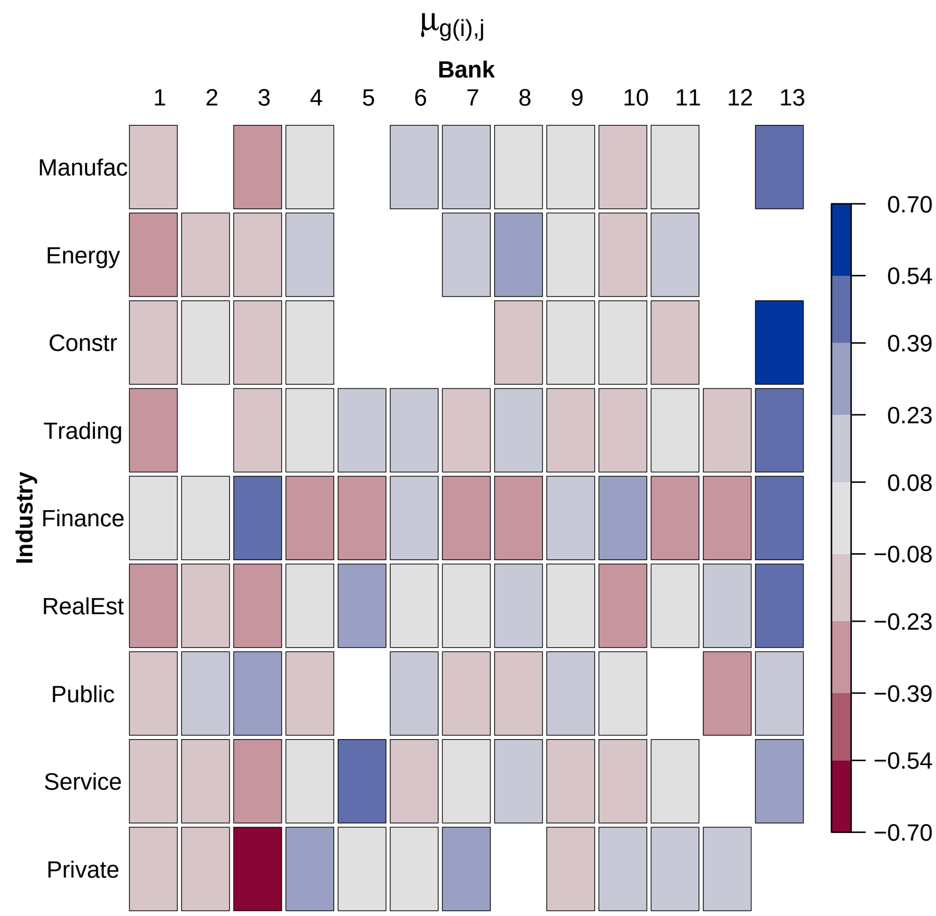
WIRTSCHAFTS UNIVERSITÄT WIEN

Industry specific means:

Industry	ν_g	$\Phi(\nu_g)$	$\Phi(\nu_g - \tau)$	$\Phi(\nu_g + \tau)$
Manufac	-2.542	55.1	17.7	151.1
Energy	-2.993	13.8	3.8	44.2
Constr	-2.448	71.8	23.8	190.9
Trading	-2.375	87.7	29.8	227.5
Finance	-3.256	5.6	1.5	19.2
RealEst	-2.474	66.8	22.6	174.7
Public	-3.330	4.3	1.1	15.1
Service	-2.517	59.2	19.8	157.0
Private	-2.296	108.4	40.0	267.4

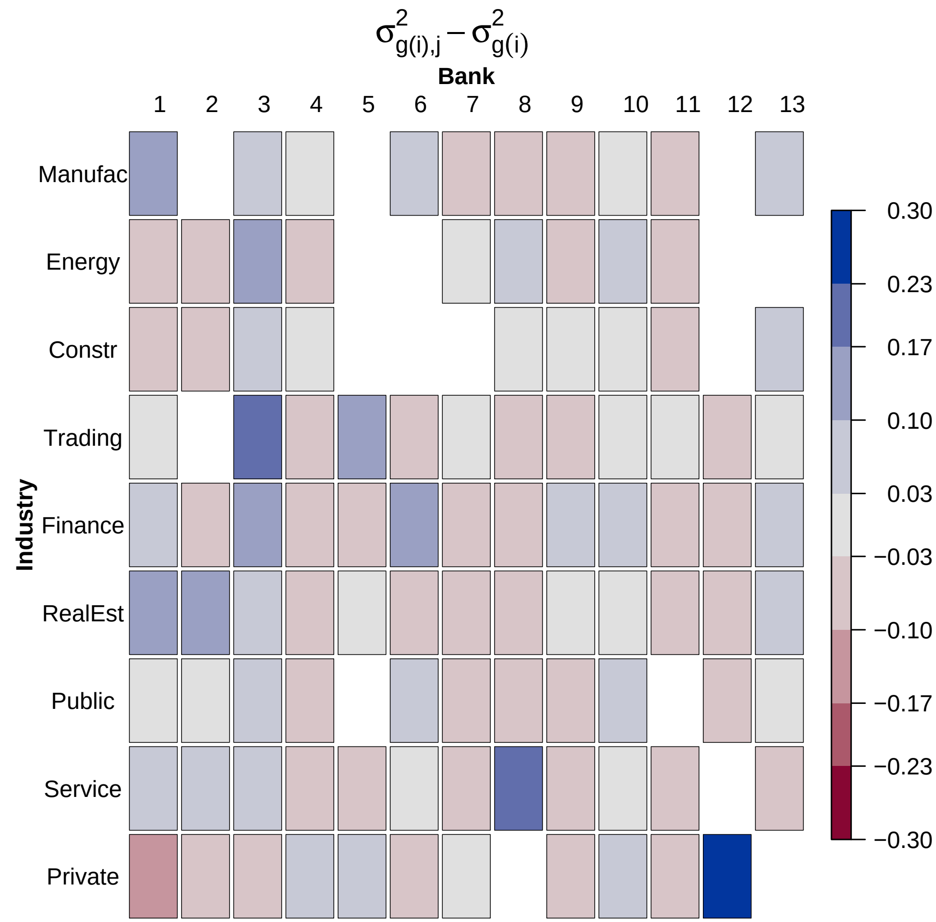
The corresponding standard deviation τ is 0.357 and the PD intervals are measured in basis points (10^{-4}).

Rating bias for rater/industry combinations:

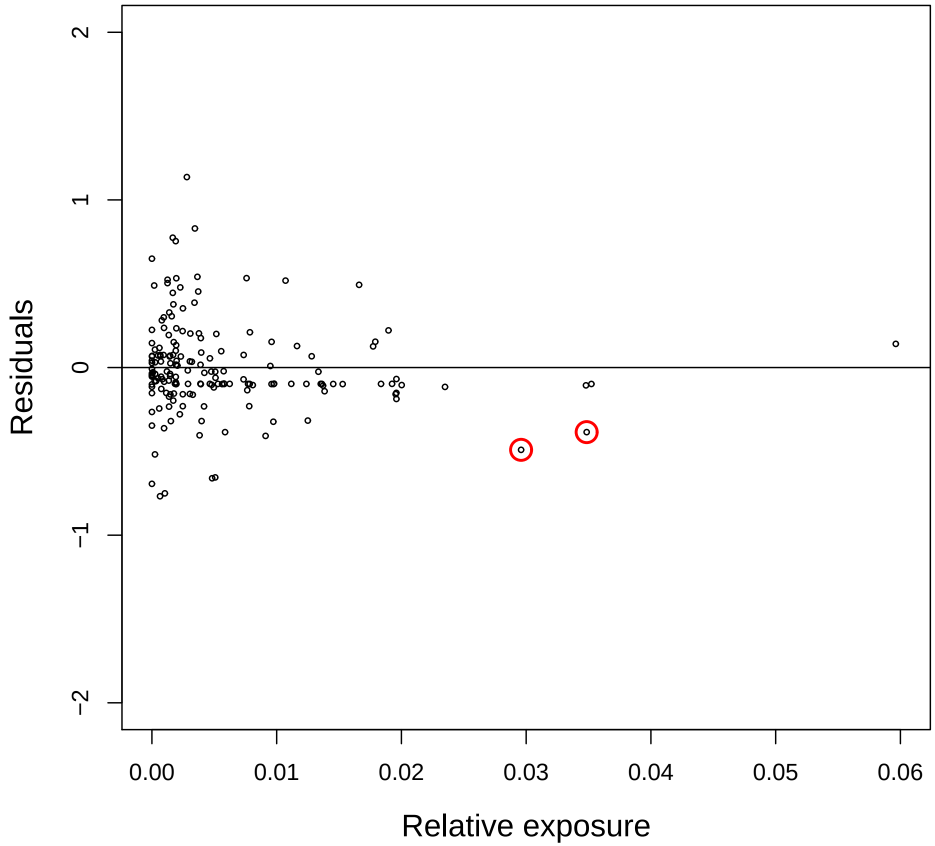


Parameters Estimates III

Rating error variance for rater/industry combinations corrected for industry effects:



Across the relative exposure:



This Bank rates two obligors with high relative exposure rather too favorably.

Motivation

- ✔ Formulate a statistical framework for a dynamic rating process:
  - ✘ Explaining the rating heterogeneity not only cross-sectionally but also longitudinally.
  - ✘ Intertemporal change of consensus ratings.
  - ✘ Effects of macroeconomic shocks – *systematic risk* (McNeil and Wendin, 2007).
  - ✘ Correlations in rating migrations (Stefanescu et al., 2009)
  - ✘ Evaluation of raters.
- ✔ Requirement: Multi-rater panel data with rating information on the same PD-scale.

- ✔ Simple Model:

$$S_{ij}(t) = S_i(t) + \mu_j + \sigma_j Z_{ij}(t),$$

  - ✘ Rating errors are independent of obligors and their characteristics (e.g., creditworthiness).
  - ✘ Parameters depend on raters only: $\mu_{ij} = \mu_j, \sigma_{ij} = \sigma_j$.
  - ✘ Latent score $S_i(t)$ is assumed to follow an AR(1) process:

$$S_i(t) = \alpha_i S_i(t-1) + \nu_i(1-\alpha_i) + \xi_i(t),$$

    where $S_i(0) = \nu_i + \xi_i(0)$, $\nu_i \sim N(\nu, \tau^2)$ and $\xi_i(t) \sim N(0, \phi^2)$.
  - ✘ AR(1) process induces auto-correlation in the latent credit worthiness of obligor i.
  - ✘ Latent scores $S_i(t)$ are uncorrelated across obligors.

Market Model

- ✔ Let $f(t)$ denote a market factor, latent score $S_i(t)$ can then be written as

$$S_i(t) = \alpha_i S_i(t-1) + \nu_i(1-\alpha_i) + \beta_{1,i} f_1(t) + \beta_{2,i} f_2(t) + \xi_i(t),$$

  - ✘ $S_i(0) = \nu_i + \xi_i(0) + \beta_{1,i} f_1(0)$.
  - ✘ $\beta_{j,i} \sim N(\mu_\beta, \sigma_\beta^2)$, $(j = 1, 2)$ denotes cross sectional correlation coefficient of scores $S_i(t)$.
  - ✘ Assume $f_1(t)$ follows AR(1) process $f_1(t) := \gamma f_1(t-1) + \zeta(t), \quad f_1(0) = \zeta(0)/\sqrt{1-\gamma^2}$, and $f_2(t) \sim N(0,1) \;\; \forall t$ to simulate market volatility shocks.

Offerings

- ✔ Assessment of rater effects (rater heterogeneity) in terms of their μ/σ^2 bias/variance characteristics.
- ✔ Estimated random effects $\hat{S}_i(t)$ can be interpreted as the consensus rating for obligor i over time.
- ✔ Market correlation structure.
  - ✘ Can the market explain the correlation structure between the obligors?

WIRTSCHAFTS UNIVERSITÄT WIEN

WIRTSCHAFTS UNIVERSITÄT WIEN

Bayesian Approach

- ✔ Both, *prior* and *likelihood* are used to derive the *posterior*
  - ✘ $p(\theta|y) \propto p(y|\theta)p(\theta)$.
- ✔ *Informative* and *non-informative* priors $p(\theta)$
  - ✘ Prior denotes beliefs about parameter before observing data.
  - ✘ Allows for expert opinions through the use of informative priors $\rightarrow$ informative prior.
  - ✘ All information resulting in a prior arose from data $\rightarrow$ noninformative prior.
- ✔ Gibbs sampling is used to derive posterior distribution.

rjags

- ✔ The R package, rjags permits a direct interface from R to the main JAGS library (Plummer, 2008).

  - ✘ `coda.samples(model, var, n.iter, thin = 5)` generates posterior sample in `mcmc.list` object
  - ✘ Function to extract random samples of the penalized deviance from `jags` model (`dic.samples(model, n.iter, thin =5, type)`).
  - ✘ includes functions for model selection (`diffdic(model1,model2)`).

- ✔ A JAGS model is defined by
  1. A model description (in a file `model.bug`).
  2. The data (a list of vectors/matrices or arrays).
  3. Optional, a set of initial values for the chain(s).

- ✔ `jags.model()` is used to create an Bayesian graphical model

```
> modeldyn <- jags.model(file = "model.bug", data = ratings,
                                inits =ratingsinit, n.chains = 4)
Compiling model graph
   Declaring variables
   Resolving undeclared variables
   Allocating nodes
   Graph Size: 23498
```

model.bug


```
model
{        f[1] <-  0;
        sf[1] <-  0;
        for (t in 2:T) {
            f[t] ~ dnorm(gamma*f[t-1], 1); sf[t] ~ dnorm(0, 1);
        }
        for (i in 1:N) {
            v[i] ~ dnorm(nu, tau);
            m[i,1] ~ dnorm(v[i] + beta[i] * f[1]+ sfbeta[i] * sf[1], sigma.i);
            for (t in 2:T) {
                m[i,t] ~ dnorm(alpha[i] * m[i,t-1] + beta[i] * f[t]  + sfbeta[i]
                                * sf[t]+ (1-alpha[i]) * v[i], sigma.i);
            }
            for (j in 1:J) {
                for (t in 1:T) {
                        y[i,j,t] ~ dnorm(m[i,t] + mu[j], sigma[j]);
                }
            }
        alpha[i] ~ dnorm(a, A); beta[i] ~ dnorm(b, B); sfbeta[i] ~ dnorm(c,C);
        }
        sigma.i ~ dgamma(1.0E-3, 1.0E-3);
        a ~ dnorm(0, 1.0E-6)T(-1.0,1.0);     A ~ dgamma(1.0E-3, 1.0E-3);
        b ~ dnorm(0, 1.0E-6);                B ~ dgamma(1.0E-3, 1.0E-3);
        c ~ dnorm(0, 1.0E-6)T(-1.0,1.0);     C ~ dgamma(1.0E-3, 1.0E-3);
        nu ~ dnorm(0, 1.0E-6);               tau ~ dgamma(1.0E-3, 1.0E-3);
        gamma ~ dnorm(0, 1.0E-6)T(-1.0,1.0);
        for (j in 1:J) {
                mu[j] ~ dnorm(0, 1.0E-6);
                sigma[j] ~ dgamma(1.0E-3,1.0E-3);
}        }
```

Output diagnostic

- Output diagnostic can be done with R package coda (Convergence Diagnosis and Output Analysis, Plummer et al., 2008).
  - coda package includes functions for convergence tests of MCMC chains.
    - `autocorr(x, lags =c() ,)` calculates the autocorrelation for the MCMC chain (`x = mcmc.obj`) at the lags `lags`.
    - `gelman.diag(x,)` calculates the scale reduction factor for a variable `x`.
    - `geweke.diag(x,)` can be be used for Geweke diagnostic.

Data

- ✔ External monthly rating data from the Big Three rating agencies Standard & Poor's, Fitch and Moody's.
- ✔ Time period: February 2007 to January 2009
- ✔ This study: 87 corporates out of the 125 corporates included in the iTraxx Europe index (Series 10)
  - ✘ iTraxx: most-liquid CDS referencing European investment-grade entities.
- ✔ Source: Historical rating announcements taken from Reuters Credit Views.

Mapping of ordinal ratings

- ✔ Rating agencies provide ratings only on a ordinal scale
  - ✘ Unclear whether PD or EL is measured.
  - ✘ Unclear which time horizon.
- ✔ Long history of default rates available.
- ✔ However, agency ratings are widely used as PD-equivalents by the industry (Erlenmaier, 2006; Tasche, 2008).
- ✔ This study: Mapping of ordinal ratings to PD provided by OENB.
- ✔ **Note:** All results of this study are conditional on used mapping procedure.
  - ✘ Quality of rating system can not be separated from mapping (in particular bias).

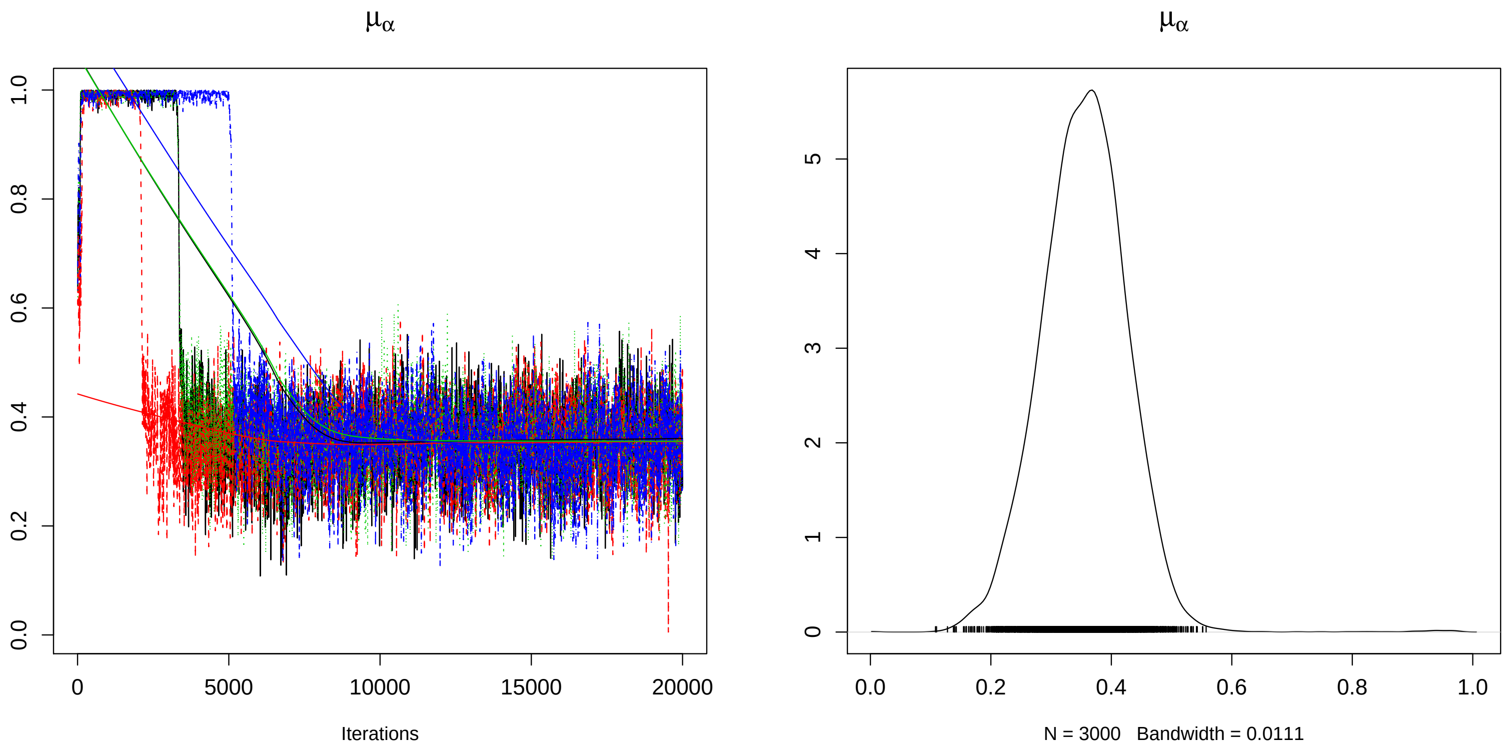
Model Selection

WIRTSCHAFTS UNIVERSITÄT WIEN

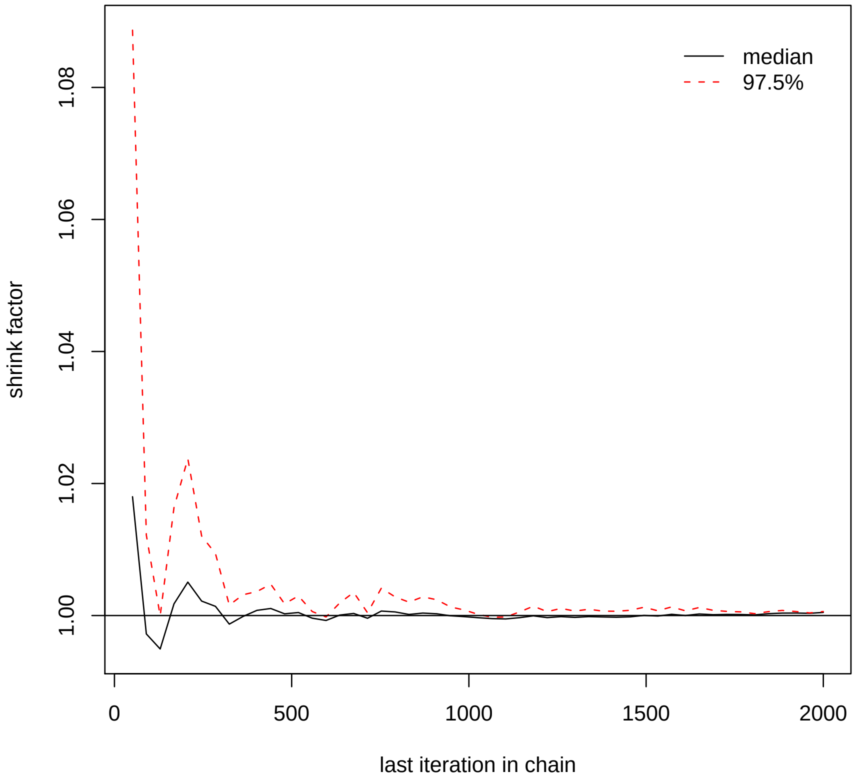
- ✔ Variety of models analyzed using deviance information criteria.
  - ✘ Simple dynamic model.
  - ✘ Dynamic market model including the EUROSTOXX50 as an exogenous market factor.
  - ✘ Dynamic market model with a market factor following an AR(1) process.
  - ✘ Dynamic market model with $f(t) = (f_1(t), f_2(t))$ with $f_1(t)$ as an AR(1) process and $f_2(t) \sim N(0, 1)$
- ✔ Dynamic market model with two market factors performed best.

Convergence Diagnosic – Trace Plot

- ✔ **Trace plot** for parameter μ_α.
  - ✘ Trace plot is plot of the iteration number against the value of drawn of parameter for each iteration
  - ✘ **Burn-in** period, about 5000 iterations



- ✔ Calculate the estimated variance of a parameter as a weighted sum of the within-chain W and between-chain variance B.
  - ✘ **Gelman Rubin** factor R is a ratio of W and B.



Rating Errors

WIRTSCHAFTSUNIVERSITÄT WIEN

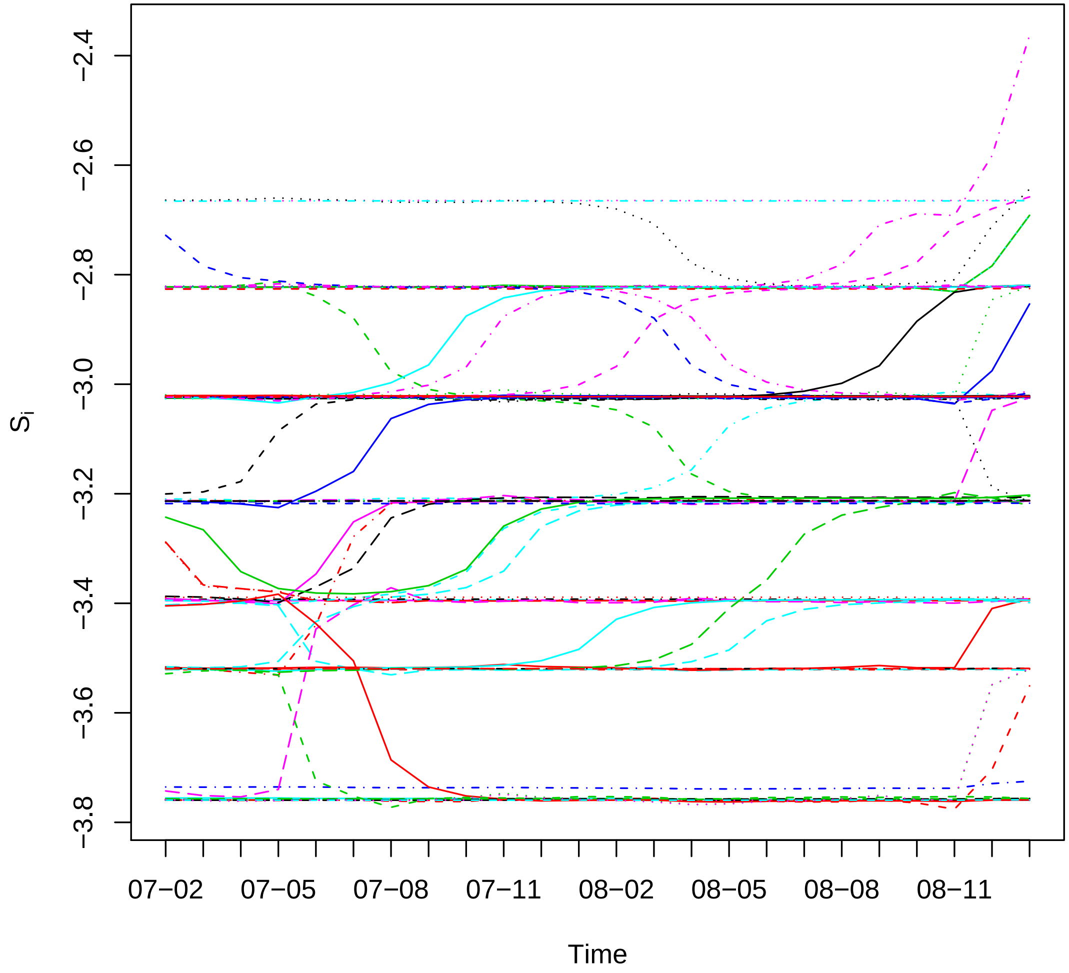
Rating bias and rating error variance for the three raters:

μ	Fitch	Moody's	S&P
mean	0.0298 (0.55 bp)	-0.1085 (-1.59 bp)	0.0787 (1.57 bp)
SD	0.0031	0.0044	0.0024

Bps relative to the grand mean ν (-3.295, in bps: 5.01).

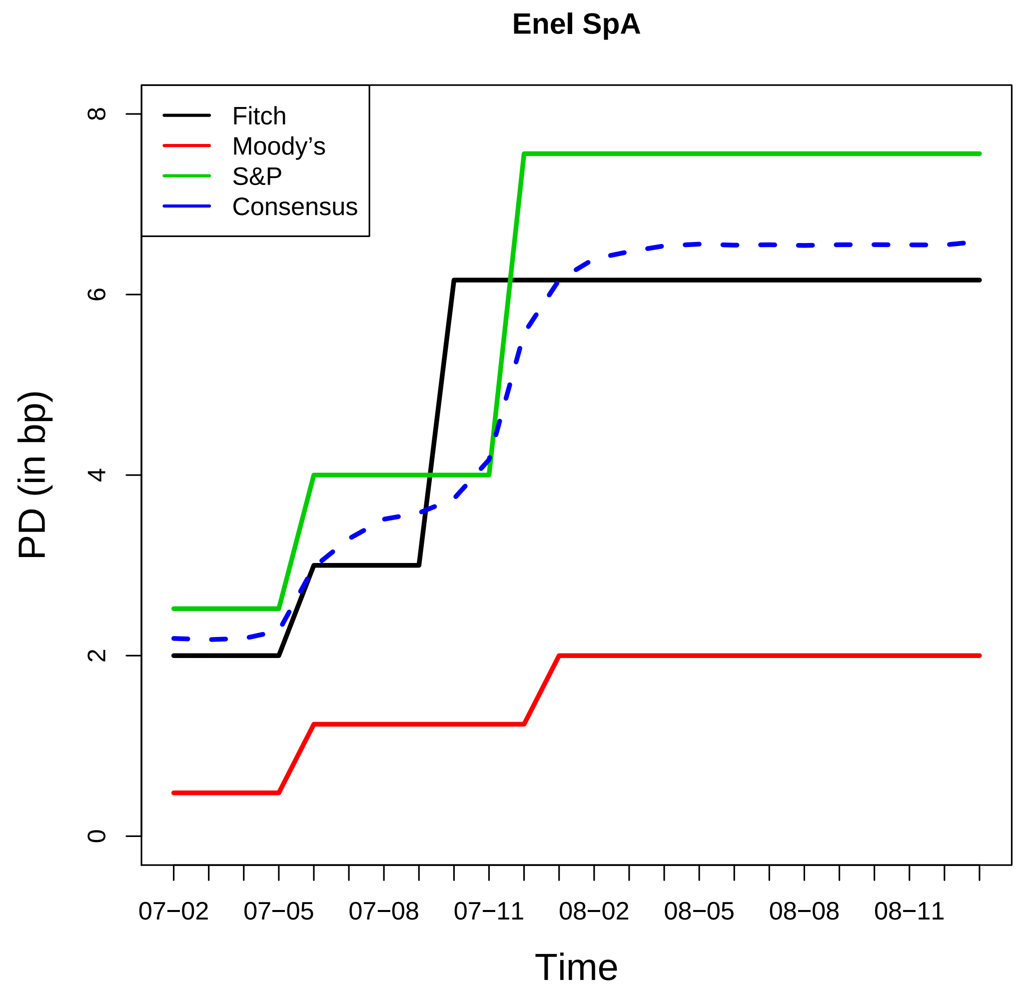
σ	Fitch	Moody's	S&P
mean	0.1443	0.1987	0.0156
SD	0.0025	0.0045	6e-04

Consensus I



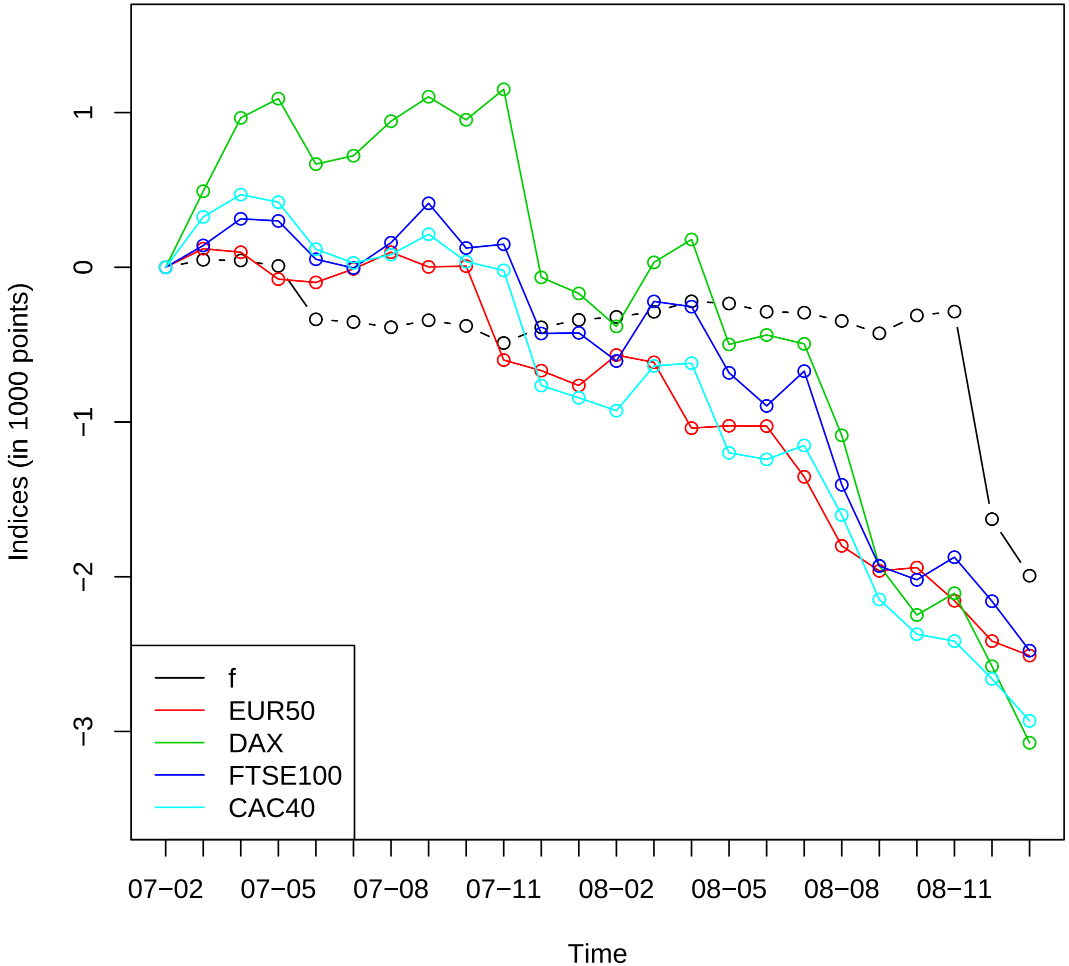
−2.4
−2.6
−2.8
−3.0
−3.2
−3.4
−3.6
−3.8
S_i
07−02
07−05
07−08
07−11
08−02
08−05
08−08
08−11
Time

Consensus II



Market Factor

WIRTSCHAFTSUNIVERSITÄT WIEN



Summary

- ✔ This methodology can be used to make fundamental statements about the heterogeneity of raters of a multi-rater panel.
- ✔ The "consensus rating" of obligors can be estimated.
- ✔ Using the consensus rating a residual analysis can be performed for single obligors as well as obligor groups of interest.
- ✔ The effect of a market factor on the consensus can be estimated.

  Requirement: The number of co-ratings within a subgroup must be sufficiently large.