A Joint Model of Failures and Credit Ratings

IAQF & Thalesians Seminar Series

Rainer Hirk joint work with Laura Vana, Stefan Pichler and Kurt Hornik

March 8, 2021

Outline

- Introduction
- Multivariate ordinal regression models
- R package **mvord**
- A joint model of ratings and failures
- Extension
- Conclusion

Outline

- Introduction
- Multivariate ordinal regression models
- R package **mvord**
- A joint model of ratings and failures
- Extension
- Conclusion

Credit risk in a nutshell

- **Credit risk** is the risk of a loss arising from a failure (or default) of a counterparty to meet its contractual obligations (e.g., McNeil et al., 2015).
- Financial intermediaries assess the credit risk of their counterparties often by using
  - statistical assessment tools, or
  - trustable third-party information.
- This is also reflected in the current global regulatory framework (e.g., Basel II (2004); Basel III (2011)).

Default-based statistical models

- Statistical credit risk models are often based on the default experience of a financial intermediary.
  - Estimation of probability of defaults (PDs) over a given time horizon (e.g., one year)
  - with a set of counterparty-specific and/or global economic variables as bankruptcy predictors
  - and a binary default indicator.
- **Advantages:**
  - Offer PDs as model outputs,
  - Flexible modeling toolbox.
- **Common problems:**
  - Defaults occur very rarely for many types of counterparties.
  - Costs of internally developing and maintaining may be high.

Credit-rating based models

- **Credit ratings** are forward-looking opinions about the creditworthiness of an obligor.
  - Credit ratings serve as a widespread alternative to internal statistical models (especially when defaults are rare).
  - Credit ratings are the most common and widely used measure of credit quality (Hilscher and Wilson, 2017).
- The three big CRAs have been intensively criticized especially in the aftermath of the 2007-2009 financial crisis.
  - For their lack of transparency and for failing to assess risk accurately.
  - CRAs react slowly to credit events and are outperformed by e.g., default-based statistical models (Lipton et al., 2012; Löffler, 2013; Kiff et al., 2013).
- **Advantages:** Easy availability and detailed classification.
- **Drawbacks:** Reliance on the correctness of external expert opinions and PDs are not provided directly.

Goal

- **Goal:** A combination of **default-based statistical models** and **credit-rating based models**
  - in order to profit from the strengths of both approaches and
  - to overcome some of the deficiencies.
- There is need for a **flexible model class** that can handle correlated ordinal and binary data:
  1. Heterogeneity in the rating methodology
  2. Heterogeneity in the covariates
  3. Unbalanced panel of firms

Outline

- Introduction
- Multivariate ordinal regression models
- R package **mvord**
- A joint model of ratings and failures
- Extension
- Conclusion

Univariate cumulative link models

- **Latent variable motivation**: the observed ordinal response Y_i is a coarser version of an underlying latent variable $\widetilde{Y}_i$:

$$\widetilde{Y}_i = \beta_0 + \mathbf{x}_i^\top \boldsymbol{\beta} + \epsilon_i, \qquad \epsilon_i \sim G, \qquad \mathbb{E}(\epsilon_i) = 0.$$

- The link between the observed variable Y_i and the latent variable $\widetilde{Y}_i$ is given by:

$$Y_i = r \quad \Leftrightarrow \quad \theta_{r-1} < \widetilde{Y}_i \leq \theta_r, \qquad r \in \{1, \dots, K\},$$

  where $-\infty = \theta_0 < \theta_1 < \cdots < \theta_{K-1} < \theta_K = \infty$ are thresholds on the latent scale.
- The cumulative link model uses the $K-1$ cumulative probabilities:

$$\mathbb{P}(Y_i \leq r | \mathbf{x}_i) = \mathbb{P}(\widetilde{Y}_i \leq \theta_r | \mathbf{x}_i) = G(\theta_r - \beta_0 - \mathbf{x}_i^\top \boldsymbol{\beta}) = \pi_{i1} + \ldots + \pi_{ir},$$

  where π_{ir} is the probability that observation i falls in the r-th category.

Extension to the multivariate setting

- $\mathbf{Y}_i = [Y_{ij}]_{j \in J_i}$ is a $(q_i \times 1)$ vector of **correlated** ordinal response variables, where
  - $i = 1, \dots, n$ is the subject index,
  - $j \in J_i \subseteq J$ denotes the outcome,
  - $q = |J|$ and $q_i = |J_i|$.
- The association between the $\mathbf{Y}_i$'s is captured by a multivariate structure imposed on the latent variables $\widetilde{\mathbf{Y}}_i$:

$$\widetilde{Y}_{ij} = \beta_{j0} + \mathbf{x}_{ij}^\top \boldsymbol{\beta}_j + \epsilon_{ij}, \qquad [\epsilon_{ij}]_{j \in J_i} = \boldsymbol{\epsilon}_i \sim F_{i,q_i}(\mathbf{0}, \boldsymbol{\Sigma}_i),$$

  where F_{i,q_i} denotes the q_i-dimensional joint distribution of the errors $\boldsymbol{\epsilon}_i$.
- For each j,

$$Y_{ij} = r \quad \Leftrightarrow \quad \theta_{j,r-1} < \widetilde{Y}_{ij} \leq \theta_{j,r}, \qquad r \in \{1, \dots, K_j\},$$

  where $-\infty = \theta_{j,0} < \theta_{j,1} < \cdots < \theta_{j,K_j-1} < \theta_{j,K_j} = \infty$ are response specific thresholds.

Choices for the multivariate distribution function

- Multivariate normal distribution $\Rightarrow$ **multivariate ordinal probit regression model**:

$$\epsilon_i \sim \mathcal{N}_{q_i}(\mathbf{0}, \mathbf{\Sigma}_i).$$

- Multivariate logistic distribution $\Rightarrow$ **multivariate ordinal logit regression model**:

$$\epsilon_i \sim \mathcal{L}_{\nu,q_i}(\mathbf{0}, \mathbf{\Sigma}_i),$$

where the multivariate logistic distribution family is constructed from a t copula with ν degrees of freedom and univariate logistic margins (O'Brien and Dunson, 2004).

Identifiability issues

- Assuming $\boldsymbol{\Sigma}_i$ to be a covariance matrix with diagonal elements $[\sigma_{ij}^2]_{j \in J_i}$, only the quantities

$$\frac{\boldsymbol{\beta}_j}{\sigma_{ij}} \quad \text{and} \quad \frac{\theta_{j,r_{ij}} - \beta_{j0}}{\sigma_{ij}} \quad \text{are identifiable.}$$

- Identifiable model parameterizations:

  1. Fixing the intercept β_{j0}, flexible thresholds $\boldsymbol{\theta}_j$ and fixing σ_{ij} $\forall j \in J_i$,
  2. Leaving the intercept β_{j0} unrestricted, fixing one threshold parameter and fixing σ_{ij},
  3. Fixing the intercept β_{j0}, fixing one threshold parameter and leaving σ_{ij} unrestricted,
  4. Leaving the intercept β_{j0} unrestricted, fixing two threshold parameters and leaving σ_{ij} unrestricted.

Pairwise likelihood estimation

- The full likelihood is approximated by a pseudo-likelihood which is constructed from lower dimensional marginal distributions.
- Let $\boldsymbol{\delta} = (\boldsymbol{\theta}, \boldsymbol{\beta}, \mathbf{P})$ denote the vector of all parameters, the **pairwise log-likelihood** function is then given by:

$$p\ell(\boldsymbol{\delta}) = \sum_{i=1}^{n} w_i \Big[\mathbb{1}_{\{q_i \geq 2\}} \sum_{\substack{k<l \\ k,l \in J_i}} \log\left(\mathbb{P}(Y_{ik} = r_{ik}, Y_{il} = r_{il})\right) + \mathbb{1}_{\{q_i = 1\}} \mathbb{1}_{\{k \in J_i\}} \log\left(\mathbb{P}(Y_{ik} = r_{ik})\right) \Big].$$

Godambe information matrix

- Under certain regularity conditions, the maximum composite likelihood estimator is consistent as $n \to \infty$ and q fixed and **asymptotically normal** with asymptotic mean $\boldsymbol{\delta}$ and covariance matrix (Varin, 2008):

$$G(\boldsymbol{\delta})^{-1} = H(\boldsymbol{\delta})^{-1} V(\boldsymbol{\delta}) H(\boldsymbol{\delta})^{-1},$$

  where
  - $G(\boldsymbol{\delta})$ denotes the **Godambe information matrix**,
  - $H(\boldsymbol{\delta})$ is the Hessian (sensitivity matrix) and
  - $V(\boldsymbol{\delta})$ is the variability matrix.

- **Standard errors** are computed using the Godambe information matrix.

- For model comparison the **composite likelihood information criterion** $\text{CLIC}(\boldsymbol{\delta}) = -2\ p\ell(\hat{\boldsymbol{\delta}}_{p\ell}) + k\ \mathrm{tr}(\widehat{V}(\boldsymbol{\delta})\widehat{H}(\boldsymbol{\delta})^{-1})$ can be used (Varin and Vidoni, 2005).

Outline

- Introduction
- Multivariate ordinal regression models
- R package **mvord**
- A joint model of ratings and failures
- Extension
- Conclusion

R package **mvord** I

- Multivariate ordinal regression models in the R package **mvord** can be fitted using the function `mvord()`.
- We offer two different **data structures**:
  - Long data format (passed by `MMO`)
  - Wide data format (passed by `MMO2`)
- **Multivariate link functions**:
  - S3 class '`mvlink`'
  - Multivariate probit and multivariate logit link
  - User is able to implement additional link functions.
- Pairwise log-likelihood is maximized by means of general purpose optimizers.

R package **mvord** II

- **Error structures:**
  - S3 class '`error_struct`'
  - Different error structures are available (general correlation or covariance, $AR(1)$, or equicorrelation).
  - Accounting for heterogeneity in the error structure among the subjects by allowing the use of subject-specific covariates in the specification of the error structure.
- **Threshold coefficients:**
  - Outcome-specific threshold coefficients
  - Constraints on the thresholds can be set by `threshold.constraints`.
  - Values can be fixed by `threshold.values`.
- **Regression coefficients:**
  - Outcome-specific regression coefficients
  - Two different designs for specifying constraints on coefficients by `coef.constraints`.
  - Values can be fixed by `coef.values`.

Methods

- Several methods are implemented for the class 'mvord'.
- These methods include summary(), print(), coef(), error_structure(), logLik(), vcov(), nobs(), terms(), model.matrix(), AIC(), BIC(), ...
- Joint probabilities can be extracted by the predict() or fitted() function:
  - type prob,
  - type cum.prob,
  - type class.
- The function marginal_predict() provides marginal predictions for the types prob, cum.prob and class.
- joint_probabilities() extracts fitted joint (cumulative) probabilities for given response categories from a fitted model.

Outline

- Introduction
- Multivariate ordinal regression models
- R package **mvord**
- A joint model of ratings and failures
- Extension
- Conclusion

Idea

- Two modeling approaches:
  - **Rating-based models** (e.g., Blume et al., 1998; Alp, 2013; Baghai et al., 2014)
  - **Default-based models** (e.g., Altman, 1968; Shumway, 2001; Tian et al., 2015)
- **Goal:** A combination of the two approaches
  - in order to profit from the strengths of both approaches and
  - to overcome some of the deficiencies.
- Advantages of such a novel estimation framework:
  - Calculate PD estimates conditional on observed external ratings.
  - Draw interesting insights from a joint distribution of ratings and defaults.

A joint model of ratings and failures

- We have a combined vector of responses $Y_i = (D_i, R_{i1}, \ldots, R_{im})$,
  - where D_i is a binary failure indicator and
  - R_{ij} is the rating observation for firm i and rater j.
- Accounting and market variables as covariates.
- Knowing the joint distribution allows to predict PDs conditional on the observed ratings in the following way:

$$\mathbb{P}(D_i = 1 | R_{i1} = r_{i1}, \ldots, R_{im} = r_{im}) = \frac{\mathbb{P}(D_i = 1, R_{i1} = r_{i1}, \ldots, R_{im} = r_{im})}{\mathbb{P}(R_{i1} = r_{i1}, \ldots, R_{im} = r_{im})}.$$

Literature on failure prediction models

- Beaver (1966) investigated the usefulness of 30 accounting ratios for failure prediction.
- Altman (1968) applied multidiscriminant analysis (Altman's Z-score).
- Ohlson (1980) performed logistic regression.
- Zmijewski (1984) applied probit regression.
- Shumway (2001) incorporated market information in a discrete time hazard model.
- Campbell et al. (2008) added new market variables and replaced the book value of the assets by their market value.
- Tian et al. (2015) performed LASSO variable selection on a set of 39 variables.

Three different sets of variables

- Five ratios of Altman's Z-score (Altman, 1968):
  - WC/TA, RE/TA, EBIT/TA, ME/LT and SALE/TA.
- Financial ratios from CHS (2008):
  - NI/MTA, TL/MTA, EXRET, IRSIZE, SIGMA, CASH/MTA, MB and PRICE.
- Ratios applied by TYG (2015):
  - LCT/TA, F/TA, NI/MTA, TL/MTA, PRICE, SIGMA and EXRET.

Data

- **Long-term issuer credit ratings** assigned by S&P, Moody's and Fitch for US companies excluding the financial and utilities sectors:
  - *Sources:* Compustat North America© Ratings File, Moody's Default & Recovery Database©, Fitch Rating Services.
- **Failure indicator:** binary indicator set to one on occurence of bankruptcy filing under Chapter 7 or Chapter 11, or default rating by CRAs in the one year-window following the rating observation;
  - *Sources:* UCLA-LoPucki Bankruptcy Research Database, Mergent FISD©.
- **Covariates:** financial ratios and market variables;
  - Pre-processing: e.g., outlier removal by winsorization, removal of missing values.
  - *Sources:* Compustat North America© Fundamentals Annual File, The Center for Research in Security Prices (CRSP).
- Period: 1985–2014

Descriptives

- 3030 firms
- 27845 firm-year observations
- In total 487 failures in the period from 1985 to 2014

	S&P	Moody's	Fitch
coverage	95.82%	58.19%	13.54%
failures	433	310	13

Model input

Model formula

```
> formula <- MMO2(failInd, SPR, Moodys, Fitch) ~ 0 + WCTA + RETA + EBITTA + MELT +
+                 SALETA
```

Function call

```
> res_SPR_Moodys_Fitch <- mvord(formula,
+                               data = data,
+                               link = mvlogit(),
+                               error.structure = cor_general(~1))
```

Altman	Failure	S&P	Moody's	Fitch
WC/TA	2.3086(0.31)***	−2.9157(0.09)***	−2.2843(0.10)***	−2.6589(0.22)***
RE/TA	1.3332(0.11)***	3.0683(0.04)***	2.9748(0.04)***	2.9355(0.06)***
EBIT/TA	9.0680(0.56)***	6.6940(0.19)***	5.7053(0.20)***	6.1007(0.37)***
ME/LT	1.0398(0.05)***	0.2290(0.01)***	0.2708(0.01)***	0.4801(0.03)***
SALE/TA	−0.2141(0.07)***	0.0878(0.02)***	0.0585(0.02)***	0.1222(0.04)***
CHS	**Failure**	**S&P**	**Moody's**	**Fitch**
NI/MTA	5.1877(0.49)***	7.7501(0.18)***	7.0623(0.22)***	6.2248(0.38)***
TL/MTA	−3.9952(0.42)***	−0.2332(0.09)***	−0.8877(0.11)***	−2.1070(0.21)***
EXRET	1.3359(0.10)***	−0.2945(0.03)***	−0.3906(0.03)***	−0.2502(0.06)***
IRSIZE	0.4228(0.05)***	1.0445(0.01)***	1.0826(0.02)***	0.8949(0.03)***
SIGMA	−0.4201(0.13)***	−1.0659(0.04)***	−1.0952(0.05)***	−0.9790(0.08)***
CASH/MTA	4.8045(0.87)***	−2.0641(0.14)***	−1.6927(0.18)***	−1.4592(0.33)***
MB	−0.4904(0.09)***	−0.5599(0.02)***	−0.5113(0.03)***	−0.4713(0.07)***
PRICE	−0.0274(0.09)	0.3111(0.03)***	−0.0454(0.03)	0.3063(0.06)***
TYG	**Failure**	**S&P**	**Moody's**	**Fitch**
LCT/TA	−1.6123(0.33)***	2.9681(0.11)***	3.0818(0.12)***	3.1102(0.21)***
F/TA	−3.5653(0.40)***	−5.4639(0.13)***	−4.9519(0.14)***	−5.0068(0.25)***
NI/MTA	4.1101(0.50)***	6.6871(0.19)***	6.3291(0.22)***	5.9433(0.37)***
TL/MTA	−3.6664(0.35)***	−1.3850(0.07)***	−2.0749(0.08)***	−2.7471(0.18)***
PRICE	0.1894(0.08)*	0.8390(0.03)***	0.5768(0.03)***	0.7193(0.05)***
SIGMA	−0.4647(0.13)***	−1.4234(0.04)***	−1.4377(0.05)***	−1.3288(0.08)***
EXRET	1.4059(0.10)***	−0.2994(0.03)***	−0.3668(0.03)***	−0.2764(0.06)***

Measuring model performance

- **Accuracy ratio** as a measure of discriminatory power

$$AR = \frac{A_R}{A_P},$$

where

  - A_R is the area between the CAP curve of the model and the random model (45 degree line)
  - A_P is the area between the CAP curve of the perfect model and the random model.

- **Weighted Brier score** as a measure of prediction accuracy

$$BS_w = \frac{\sum_{i=1}^{n} w_i (p_i - d_i)^2}{\sum_{i=1}^{n} w_i},$$

where

  - w_i are weights for each observation pair,
  - p_i is the predicted PD,
  - d_i is 1 for failed firms and 0 otherwise.

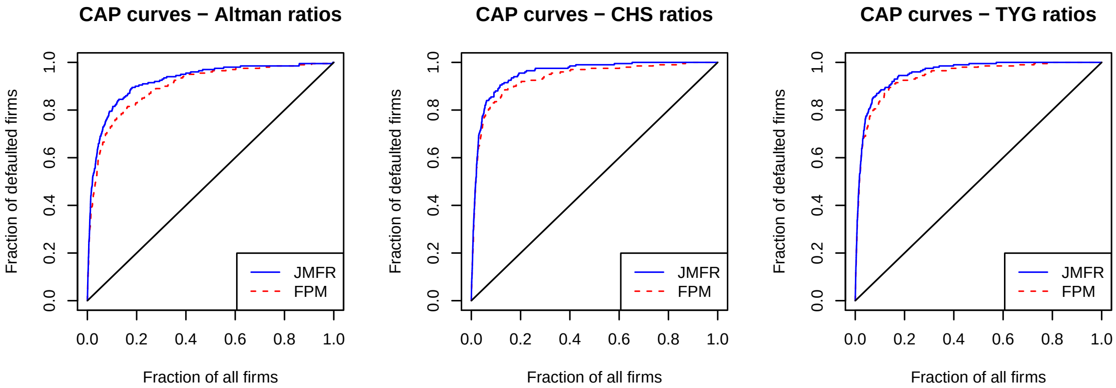
Model comparison

- **Training sample**: randomly selected 60% of the firms
- **Test sample**: remaining 40% of the firms

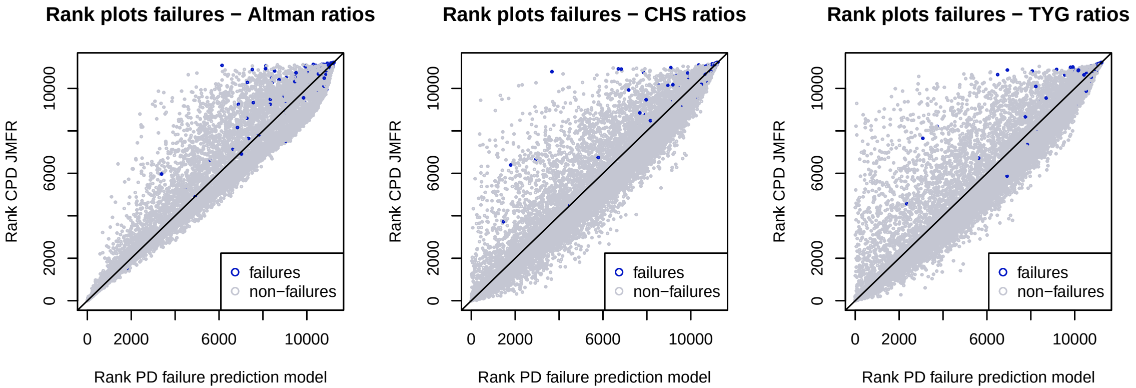
Model	Altman			CHS			TYG		
	AR	Brier	w. Brier	AR	Brier	w. Brier	AR	Brier	w. Brier
FPM[1]	0.8112	0.0141	0.3505	0.8828	0.0128	0.3027	0.8968	0.0120	0.2852
JMFR[2] S+M+F	0.8587	0.0129	0.3065	0.9232	0.0121	0.2758	0.9257	0.0113	0.2602
JMFR S	0.8539	0.0132	0.3120	0.9217	0.0122	0.2830	0.9212	0.0114	0.2645
JMFR M	0.8157	0.0143	0.3615	0.8813	0.0128	0.3043	0.8989	0.0121	0.2867
JMFR F	0.8111	0.0141	0.3517	0.8829	0.0128	0.3027	0.8972	0.0120	0.2851
JMFR S + M	0.8577	0.0128	0.3040	0.9235	0.0120	0.2763	0.9254	0.0113	0.2586
JMFR S + F	0.8544	0.0132	0.3145	0.9222	0.0122	0.2822	0.9220	0.0115	0.2657
JMFR M + F	0.8157	0.0143	0.3615	0.8813	0.0128	0.3043	0.8989	0.0121	0.2867

[1] FPM abbreviates failure prediction model.
[2] JMFR abbreviates joint model of failures and ratings.

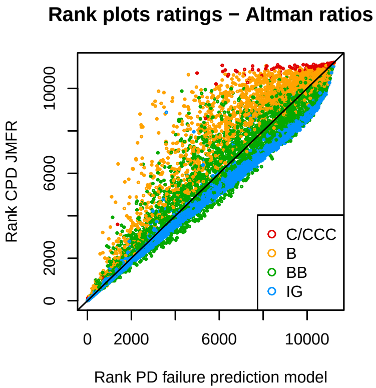
CAP curves

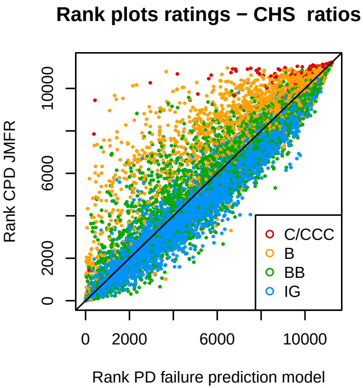

CAP curves − Altman ratios
CAP curves − CHS ratios
CAP curves − TYG ratios
Fraction of defaulted firms
Fraction of all firms
0.0
0.2
0.4
0.6
0.8
1.0
JMFR
FPM

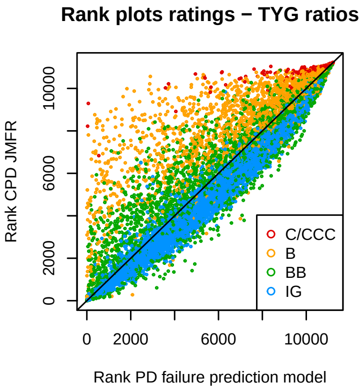
Rank plots JMFR vs. FPM



Rank plots JMFR vs. FPM - ratings

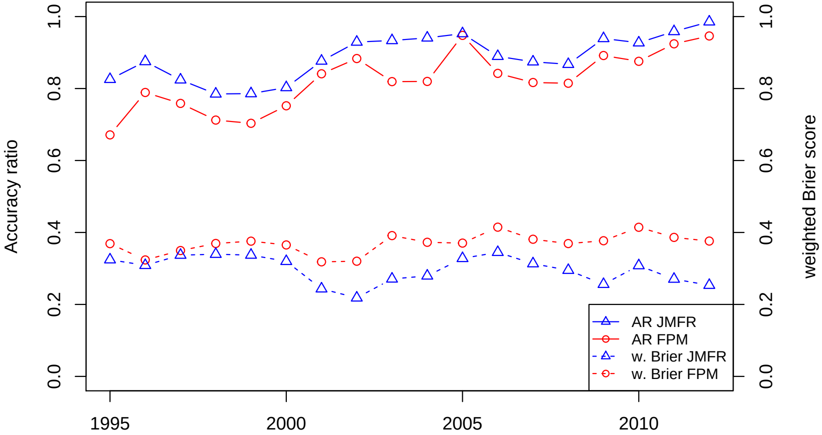






Rolling windows

- 10 years in-sample training
- 2 year out-of-sample predictions



Outline

The mvordflex model I

- Let $Y^j_{i,t}$ denote an ordinal observation,
  - where $i = 1, \dots, n$ denotes the subject index,
  - $t = t_1, t_2, \dots T$ are equidistant time points and
  - $j \in J_i$, $q = |J|$ and $q_i = |J_i|$ denotes the cardinality of the sets J and J_i.
- We assume the ordinal observation $Y^j_{i,t}$ to be a coarser version of a continuous latent variable

$$\tilde{Y}^j_{i,t} = (\mathbf{x}^j_{i,t})^\top \beta^j_t + \epsilon^j_{i,t}$$

connected by a vector of suitable threshold parameters $\boldsymbol{\theta}$:

$$Y^j_{i,t} = r^j_{i,t} \Leftrightarrow \theta_{r^j_{i,t}-1} < \tilde{Y}^j_{i,t} \leq \theta_{r^j_{i,t}}, \quad r^j_{i,t} \in \{1, \dots, K_j\},$$

where $r^j_{i,t}$ is one of the K_j ordered categories. For each outcome j and time point t, we have the following restriction on $\boldsymbol{\theta}^j_t$: $-\infty \equiv \theta^j_{t,0} < \theta^j_{t,1} < \cdots < \theta^j_{t,K_j-1} \equiv \infty$.

The mvordflex model II

- For $\mathbf{X}^*_{i,t} = (\mathbf{I}_q \otimes \mathbf{x}^\top_{i,t}) = \begin{pmatrix} \mathbf{x}^\top_{i,t} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x}^\top_{i,t} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x}^\top_{i,t} \end{pmatrix}$ with $\mathbf{x}_{i,t}$ being the p-dimensional vector of covariates and
- β^*_t a $p \cdot q$-dimensional vector $\beta^*_t = ((\beta^1_t)^\top, (\beta^2_t)^\top, \ldots, (\beta^q_t)^\top)^\top$
- we obtain:

$$\tilde{\mathbf{Y}}_{i,t} = \mathbf{X}^*_{i,t}\beta^*_t + \epsilon_{i,t},$$

with

$$\epsilon_{i,t} = \Psi\epsilon_{i,t-1} + \mathbf{u}_{i,t}$$
$$u_{i,t} \sim N_q(\mathbf{0}, \mathbf{\Sigma})$$

The mvordflex model III

- Inter-rater dependence

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,q} \\ \rho_{1,2} & \ddots & \ddots & \rho_{2,q} \\ \vdots & \ddots & \ddots & \vdots \\ \rho_{1,q} & \rho_{2,q} & \cdots & 1 \end{pmatrix}$$

- time dependence for each rater

$$\boldsymbol{\Psi} = \text{diag}(\rho_1, \rho_2, \ldots, \rho_q)$$

Outlook

- **mvordflex**: Time varying model for (multiple) raters and failures
  - set up modeling framework with composite likelihood methods
  - performed a simulation study to investigate the quality of the estimates in different parameter settings
  - find a high time persistence in ratings
- Include time-varying coefficients
- Compare conditional PDs to previous models
- Rating transition probabilities
- Sector-specific error structures

Outline

- Introduction
- Multivariate ordinal regression models
- R package **mvord**
- A joint model of ratings and failures
- Extension
- Conclusion

Conclusion

- We propose a joint modeling framework, where
  - binary failure information and
  - credit ratings or expert opinions can be included.
- The proposed multivariate framework
  - is able to account for missing observations in the response variables and
  - offers PD estimates conditional on the observed ratings at the beginning of the period.
- We find that adding rating information in a failure prediction models gives an improvement in the predictive performance and discriminatory power.

References

Edward I. Altman. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23 (4): 589–609, 1968.

John Y Campbell, Jens Hilscher, and Jan Szilagyi. In search of distress risk. *The Journal of Finance*, 63 (6): 2899–2939, 2008.

Rainer Hirk, Kurt Hornik, and Laura Vana. **mvord**: An R package for fitting multivariate ordinal regression models. *Journal of Statistical Software*, 93 (4):1–41, 2020

Rainer Hirk, Laura Vana, Kurt Hornik, and Stefan Pichler. A joint model of failures and credit ratings. *Journal of Credit Risk*, 2020.

A. J. McNeil, R., Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princeton, NJ, USA, 2015.

Shaonan Tian, Yan Yu, and Hui Guo. Variable selection and corporate bankruptcy forecasts. *Journal of Banking & Finance*, 52: 89–100, 2015.

Thank you for your attention!

Rainer Hirk
rhirk@wu.ac.at

Institute for Statistics and Mathematics
WU Vienna University of Economics and Business
Welthandelsplatz 1
1020 Vienna

References I

Basel II: International Convergence of Capital Measurement and Capital Standards: a Revised Framework. Online publication, 2004. URL `http://www.bis.org/publ/bcbs107.htm`.

Basel III: A global regulatory framework for more resilient banks and banking systems. Online publication, 2011. URL `https://www.bis.org/publ/bcbs189.htm`.

Aysun Alp. Structural shifts in credit rating standards. *The Journal of Finance*, 68(6):2435–2470, 2013. doi: 10.1111/jofi.12070.

Edward I. Altman. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4):589–609, 1968. doi: 10.2307/2978933.

Ramin P. Baghai, Henri Servaes, and Ane Tamayo. Have rating agencies become more conservative? Implications for capital structure and debt pricing. *The Journal of Finance*, 69(5):1961–2005, 2014.

William H. Beaver. Financial ratios as predictors of failure. *Journal of Accounting Research*, 4:71–111, 1966. ISSN 00218456, 1475679X. URL `http://www.jstor.org/stable/2490171`.

Marshall E. Blume, Felix Lim, and A. Craig Mackinlay. The declining credit quality of u.s. corporate debt: Myth or reality? *The Journal of Finance*, 53(4):1389–1413, 1998. doi: 10.1111/0022-1082.00057.

John Y Campbell, Jens Hilscher, and Jan Szilagyi. In search of distress risk. *The Journal of Finance*, 63(6):2899–2939, 2008.

Jens Hilscher and Mungo Wilson. Credit ratings and credit risk: Is one measure enough? *Management Science*, 63(10): 3414–3437, 2017. doi: 10.1287/mnsc.2016.2514.

References II

John Kiff, Michael Kisser, and Miss Liliana Schumacher. *Rating through-the-cycle: what does the concept imply for rating stability and accuracy?* Number 13-64. International Monetary Fund, 2013.

Alexander Lipton, Andrew Rennie, and Edward I. Altman. Default recovery rates and lgd in credit risk modelling and practice, 09 2012. URL `http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199546787.001.0001/oxfordhb-9780199546787-e-3`.

Gunter Löffler. Can rating agencies look through the cycle? *Review of Quantitative Finance and Accounting*, 40(4): 623–646, 2013.

A. J. McNeil, R., Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princeton, NJ, USA, 2015.

Sean M. O'Brien and David B. Dunson. Bayesian multivariate logistic regression. *Biometrics*, 60(3):739–746, 2004. doi: 10.1111/j.0006-341X.2004.00224.x.

James A Ohlson. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research*, pages 109–131, 1980.

Tyler Shumway. Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business*, 74(1): 101–124, 2001. doi: 10.1086/209665.

Shaonan Tian, Yan Yu, and Hui Guo. Variable selection and corporate bankruptcy forecasts. *Journal of Banking & Finance*, 52:89–100, 2015. ISSN 0378-4266. doi: 10.1016/j.jbankfin.2014.12.003.

Cristiano Varin. On composite marginal likelihoods. *AStA Advances in Statistical Analysis*, 92(1):1, Feb 2008. ISSN 1863-818X. doi: 10.1007/s10182-008-0060-7.

References III

Cristiano Varin and Paolo Vidoni. A note on composite likelihood inference and model selection. *Biometrika*, 92(3): 519–528, 2005. doi: 10.1093/biomet/92.3.519.

Mark E Zmijewski. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting research*, pages 59–82, 1984.