Multivariate Ordinal Regression Models: An Analysis of Corporate Credit Ratings

Hirk, Rainer; Hornik, Kurt; Vana Gür, Laura

Published: 01/01/2017

Document Version
Publisher's PDF, also known as Version of record

Multivariate Ordinal Regression Models: An Analysis of Corporate Credit Ratings

Rainer Hirk, Kurt Hornik, Laura Vana

Research Report Series

Report 132, January 2017

Institute for Statistics and Mathematics

http://statmath.wu.ac.at/

Multivariate Ordinal Regression Models: An Analysis of Corporate Credit Ratings

Rainer Hirk[1] Kurt Hornik[1] Laura Vana *[1]

[1] WU Vienna University of Economics and Business

This version: January 23, 2017

Abstract

Correlated ordinal data typically arise from multiple measurements on a collection of subjects. Motivated by an application in credit risk, where multiple credit rating agencies assess the creditworthiness of a firm on an ordinal scale, we consider multivariate ordinal models with a latent variable specification and correlated error terms. Two different link functions are employed, by assuming a multivariate normal and a multivariate logistic distribution for the latent variables underlying the ordinal outcomes. Composite likelihood methods, more specifically the pairwise and tripletwise likelihood approach, are applied for estimating the model parameters. We investigate how sensitive the pairwise likelihood estimates are to the number of subjects and to the presence of observations missing completely at random, and find that these estimates are robust for both link functions and reasonable sample size. The empirical application consists of an analysis of corporate credit ratings from the big three credit rating agencies (Standard & Poor's, Moody's and Fitch). Firm-level and stock price data for publicly traded US companies as well as an incomplete panel of issuer credit ratings are collected and analyzed to illustrate the proposed framework.

Keywords: composite likelihood, credit ratings, financial ratios, latent variable models, multivariate ordered probit, multivariate ordered logit

*Correponding author: Laura Vana, Department of Finance, Accounting and Statistics, WU Vienna University of Economics and Business, Welthandelsplatz 1, D4/4, Vienna, Austria, Tel.: +43-1-31336-5032, E-mail: laura.vana@wu.ac.at.

1 Introduction

The analysis of univariate or multivariate ordinal outcomes is an important task in various fields of research from social sciences to medical and clinical research. A typical setting where correlated ordinal outcomes arise naturally is when several raters assign different ratings on a collection of subjects. In the financial markets literature ordinal data often appear in the form of credit ratings (e.g., Cantor and Packer, 1997; Blume et al., 1998; Bongaerts et al., 2012; Becker and Milbourn, 2011; Alp, 2013). Credit ratings are ordinal rankings of credit risk, i.e., the risk of a firm not being able to meet its financial obligations, and are typically produced by credit rating agencies (CRAs). Especially in the US, CRAs like Standard and Poor's (S&P), Moody's and Fitch play a significant role in financial markets, with their credit ratings being one of the most common and widely used sources of information about credit quality.

The CRAs provide in their issuer ratings a forward-looking opinion on the total creditworthiness of a firm. In evaluating credit quality, quantitative and qualitative criteria are employed. The quantitative analysis relies mainly on the assessment of market conditions and on financial analysis. Key financial ratios, built from market information and financial statements, are used to evaluate several aspects of a firm's performance (according to Puccia et al., 2013, such aspects are profitability, leverage, cash-flow adequacy, liquidity, and financial flexibility). In credit risk modeling, the literature on credit ratings so far usually considered separate models for one or more CRAs. For example, Blume et al. (1998) as well as Alp (2013) use ordinal regression models with financial ratios as explanatory variables to obtain insights into the rating behavior of S&P.

In general, the ratings from the big three CRAs do not always coincide and they sometimes differ by several rating notches due to multiple reasons. First, S&P and Fitch use different rating scales compared to Moody's. Second, S&P and Fitch consider probabilities of default as the key measure of creditworthiness, while Moody's ratings also incorporate information about recovery rates in case of default. Third, given the fact that the rating and estimation methodology of the CRAs is not completely disclosed, there is ambiguity about whether the CRAs give different importance to different covariates in their analysis. In view of these facts, a multivariate analysis, where credit ratings are considered as dependent variables and firm-level and market information as covariates, provides useful insights into heterogeneity among different raters and into determinants of such credit ratings.

To motivate this study we focus on a data set of US corporates over the period 2000–2013 for which at least one corporate credit rating from the big three CRAs is available. For this purpose we propose the use of multivariate ordered probit and logit regression models. The proposed models incorporate non-standard features, such as different threshold parameters and different regression coefficients for each outcome variable to accommodate for the different scales and methodologies of the CRAs. Observations missing completely at random will also be handled by the model, as not all firms are rated by all CRAs at the same point in time. Aside from the inferred relationship between the outcomes and various relevant covariates based on the regression coefficients, multivariate ordinal regression models allow inference on the agreement between the different raters. Using the latent variable specification, where each ordinal variable represents a discretized version of an underlying latent continuous random variable, association can be measured by the correlation between these latent variables. The complexity of the model can further be increased by letting the correlation parameters further depend on covariates. In our application we only consider business sectors as relevant covariates for the correlation structure.

Estimation of the multivariate ordered probit and logit models is performed using composite likelihood methods. These methods reduce the computational burden by replacing the full likelihood by a product of lower-dimensional component likelihoods. For the logit link we employ the multivariate logistic distribution of O'Brien and Dunson (2004), which is approximated by multivariate t distribution with marginal logistic distributions. The use of the multivariate t distribution allows for a flexible correlation matrix.

While multivariate linear models have been extensively researched and applied, multivariate modeling of discrete or ordinal outcomes is more difficult, owing to the lack of analytical tractability and computational convenience. However, many advances have been made in the last two decades. An overview of statistical modeling of ordinal data is provided by Greene and Hensher (2010) and Agresti (2010). The main approaches to formulate multivariate ordinal models include: (i) modeling the mean levels and the association between responses at a population level by specifying marginal distributions; such marginal models are estimated using generalized estimating equations. (ii) Under the latent variable specification, joint distribution functions are assumed for the latent variables underlying the ordinal outcomes. Estimation of multivariate ordinal models in the presence of covariates can be performed using Bayesian and frequentist techniques. Chib and Greenberg (1998) and Chen and Dey (2000) were among the first to perform a fully Bayesian analysis of multivariate binary and ordinal outcomes, respectively, and to develop several Metropolis Hastings algorithms to simulate the posterior distributions of the parameters of interest. Difficulties in Bayesian inference arise due to the fact that absolute scale is not identifiable in ordinal models. In this case, the covariance matrix of the multiple outcomes is often restricted to be a correlation matrix which makes the sampling of the correlation parameters non-standard. Moreover, threshold parameters are typically highly correlated with the latent responses. Bayesian semi- or non-parametric techniques can be employed if normality of the latent variables is assumed to be a too restrictive assumption (e.g., Kim and Ratchford, 2013; DeYoreo and Kottas, 2014). Nonetheless, research into these techniques is still on-going.

Frequentist estimation techniques include maximum likelihood (e.g., Scott and Kanaroglou, 2002; Nooraee et al., 2016), which is usually feasible for a small number of outcomes. If the multivariate model for the latent outcomes is formulated as a mixed effects model with correlated random effects, Laplace or Gauss-Hermite approximations, as well as EM algorithms can be applied. EM algorithms treat the random effects as missing observations can be employed to estimate the model parameters (Grigorova et al. 2013 extended the EM algorithm for the univariate case of Kawakatsu and Largey 2009 to the multivariate case). However, we experienced convergence problems in our application. Alternatively, estimation using maximum simulated likelihood has been proposed (e.g., Bhat and Srinivasan, 2005), which uses (quasi) Monte Carlo methods to approximate the integrals in the likelihood function. This method has been reported to be unstable and to suffer from convergence issues as the dimension of the outcomes increases (a simulation study is provided by Bhat et al., 2010). An estimation method which has managed to overcome most of the difficulties faced by other techniques is the composite likelihood method, which can easily be employed for higher number of ordinal outcomes measured on a cross-section (e.g., Bhat et al., 2010; Pagui et al., 2015). In addition, the composite likelihood estimator has satisfactory asymptotic properties (a comprehensive overview on the theory, efficiency and robustness of this estimator is provided by Varin et al., 2011).

This paper is organized as follows: Section 2 provides an overview of multivariate ordinal regression models, including model formulation, link functions and identifiability issues. Estimation is discussed in Section 3. In Section 4 we set-up an extensive simulation study and investigate how different aspects and characteristics of the data influence the accuracy of the estimates. The multiple credit ratings data set is analyzed in Section 5. Section 6 concludes.

2 Model

Several models can be employed for ordinal data analysis with *cumulative link models* being the most popular ones. A cumulative link model can be motivated by assuming that the observed ordinal variable Y is a coarser (categorized) version of a latent continuous variable $\widetilde{Y}$.

Suppose one has a (possibly unbalanced) panel of firms observed repeatedly over T years with a total of n firm-year observations. Moreover, suppose each firm h in year t is assigned a rating on an ordinal scale by CRAs indexed by $j \in J_{ht}$, where J_{ht} is a non-empty subset from the set of all J available raters[1]. The number of available outcomes for firm h in year t is given by q_{ht}. Let Y_{htj} denote the rating assigned by rater j to firm h in year t. The observable categorical outcome Y_{htj} with K_j possible ordered categories and the unobservable latent variable $\widetilde{Y}_{htj}$ are connected by:

$$Y_{htj} = r_{htj} \quad \text{if } \theta_{j,r_{htj}-1} < \widetilde{Y}_{htj} \leq \theta_{j,r_{htj}}, \qquad r_{htj} \in \{1, \dots, K_j\},$$

where $\boldsymbol{\theta}_{j,\cdot}$ is a vector of suitable threshold parameters for outcome j with the following restriction: $-\infty \equiv \theta_{j,0} < \theta_{j,1} < \cdots < \theta_{j,K_j} \equiv \infty$. We allow the thresholds to vary across outcomes to account for differences in the rating behavior of each rater.

Given an $n \times p$ covariate matrix X, where each row $\boldsymbol{x}_{ht}$ is a p-dimensional vector of covariates for firm h in year t, we assume the following linear model:

$$\widetilde{Y}_{htj} = \beta_{0j} + \alpha_{tj} + \boldsymbol{x}_{ht}^\top \boldsymbol{\beta}_j + \epsilon_{htj}, \qquad [\epsilon_{htj}]_{j \in J_{ht}} = \boldsymbol{\epsilon_{ht}} \sim F_{ht,q_{ht}}, \tag{1}$$

where β_{0j} is a constant term, α_{tj} is an intercept for year t and rater j, $\boldsymbol{\beta}_j$ is a vector of slope coefficients corresponding to outcome j[2] and ϵ_{htj} is a mean zero error term distributed according to a q_{ht}-dimensional distribution function $F_{ht,q_{ht}}$. In the case of observations missing completely at random ($q_{ht} < J$), $F_{ht,q_{ht}}$ is the marginal distribution corresponding to the distribution function of the errors in case all J ratings had been observed.

[1] For example, if firm h in year t is rated by raters one and three out of a total of three raters ($J = 3$), one has the set $J_{ht} = \{1, 3\}$.

[2] Note that this setting easily accommodates the use of different covariates for each outcome, by restricting a-priori some of the slope coefficients to zero.

The year intercepts should capture stringency or loosening of the rating standards of each CRA relative to a baseline year, in our case the first year in the sample (like in Blume et al., 1998; Alp, 2013; Baghai et al., 2014). The different $\boldsymbol{\beta}_j$'s are able to account for heterogeneity in the rating methodology of the raters. We assume that errors are independent across firms and years with distribution function $F_{ht,q_{ht}}$ and orthogonal to the covariates. Longitudinal correlation structures in the errors, like an auto-regressive model of order one, could capture the effect of the business cycle on the creditworthiness of a firm. We, however, do not incorporate such structure into the errors, motivated by the practice of the CRAs of rating "through the cycle" i.e., the ratings should not respond to temporary fluctuations in credit quality caused by economic cycle effects (Puccia et al., 2013).

In order to simplify notation, the $n \times (T-1)$ matrix of year dummies D will be incorporated together with the covariates into a new matrix $\widetilde{X} = (D\ X)$ and the vector $\widetilde{\boldsymbol{\beta}}_j = (\boldsymbol{\alpha}_j^\top, \boldsymbol{\beta}_j^\top)^\top$ will contain the $T-1$ year intercepts $\boldsymbol{\alpha}_j$ and the vector of slope coefficients $\boldsymbol{\beta}_j$. Using this notation, the index ht for each firm-year observation is replaced by $i = \{1, \dots, n\}$, and we call each firm-year observation hereafter a subject. Thus, model (1) becomes:

$$\widetilde{Y}_{ij} = \beta_{0j} + \widetilde{\boldsymbol{x}}_i^\top \widetilde{\boldsymbol{\beta}}_j + \epsilon_{ij}, \qquad [\epsilon_{ij}]_{j \in J_i} = \boldsymbol{\epsilon_i} \sim F_{i,q_i}. \tag{2}$$

Link functions The distribution functions we consider for the error terms are the multivariate normal and logistic distributions, where the corresponding models for the observed variable Y_{ij} are the cumulative probit and the cumulative logit link models.

The probit link arises if the error terms in model (1) are assumed to follow a multivariate normal distribution: $\boldsymbol{\epsilon_i} \sim MVN_{q_i}(\mathbf{0}, \boldsymbol{\Sigma_i})$. In defining a multivariate logistic distribution, we follow the lines of O'Brien and Dunson (2004), who proposed an approximate multivariate logistic density, which they employ for performing posterior inference in a Bayesian multivariate logistic regression. Their approach has been adopted by Nooraee et al. (2016) in a frequentist setting, who use maximum likelihood for estimating a multivariate ordinal model for longitudinal data. The proposed multivariate logistic density with location $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ for J dimensions is:

$$\mathcal{L}_J(\boldsymbol{z}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \mathcal{T}_{J,\nu}(\boldsymbol{t}|\mathbf{0}, \boldsymbol{R}) \prod_{j=1}^{J} \frac{\mathcal{L}(z_j|\mu_j, s_j)}{\mathcal{T}_\nu(t_j|0,1)}, \tag{3}$$

where $[s_j^2]_{j \in \{1,\dots,J\}}$ are the diagonal elements of $\boldsymbol{\Sigma}$, $\boldsymbol{R}$ is the correlation matrix corresponding to $\boldsymbol{\Sigma}$ and $t_j = \mathrm{T}_\nu^{-1}(\exp((z_j - \mu_j)/s_j)/(1 + \exp((z_j - \mu_j)/s_j)))$ with T_ν^{-1} being the inverse univariate standard t distribution function with ν degrees of freedom; $\mathcal{T}_{J,\nu}(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the multivariate t density with ν degrees of freedom, location $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$, $\mathcal{L}(\cdot|\mu, s)$ denotes the univariate logistic density with location μ and scale s and $\mathcal{T}_\nu(\cdot|\mu, \sigma)$ denotes the univariate t density with location μ, scale σ and ν degrees of freedom. As previously shown by Albert and Chib (1993), the univariate logistic density with location parameter μ and scale s is well approximated by a t distribution. The two densities are approximately equivalent when setting $\sigma^2 = \tilde{\sigma}^2 \equiv s^2\pi^2(\nu-2)/3\nu$ and $\nu = \tilde{\nu} \equiv 8$ (cf. Nooraee et al., 2016). The same property holds for the proposed multivariate density such that $\mathcal{L}_J(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \approx \mathcal{T}_{J,\tilde{\nu}}\left(\cdot|\boldsymbol{\mu}, \frac{\pi^2(\tilde{\nu}-2)}{3\tilde{\nu}}\boldsymbol{\Sigma}\right)$.

Gumbel (1961) was the first to propose a bivariate logistic distribution which was later extended to the multivariate case by Malik and Abraham (1973). This multivariate distribution has only one parameter to represent the dependence between all outcomes. The main advantages of using the multivariate logistic distribution in Equation (3) are i) it allows for a flexible unconstrained correlation structure between the underlying latent variables $\widetilde{Y}$ and ii) the regression coefficients can be interpreted in terms of log odds ratios.

Identifiability It is well known that in ordinal models the absolute location and the absolute scale of the underlying latent variable are not identifiable (see for example Chib and Greenberg, 1998). Assuming that $\boldsymbol{\Sigma}_i$ is the full covariance matrix of the errors $\boldsymbol{\epsilon}_i$ with diagonal elements $[\sigma_{ij}^2]_{j\in J_i}$, in model (2) only the quantities $\widetilde{\boldsymbol{\beta}}_j/\sigma_{ij}$ and $(\theta_{j,r_{ij}} - \beta_{0j})/\sigma_{ij}$ are identifiable. As such, typical constraints on the parameters are, for all j:

- fixing β_{0j} (e.g., to zero), using flexible thresholds $\boldsymbol{\theta}_{j,\cdot}$ and fixing σ_{ij} (e.g., to unity);
- leaving β_{0j} unrestricted, fixing one threshold parameter (e.g., $\theta_{j,1} = 0$), fixing σ_{ij} (e.g., to unity);
- leaving β_{0j} unrestricted, fixing two threshold parameters (e.g., $\theta_{j,1} = 0$ and $\theta_{j,K_j-1} = 1$), leaving σ_{ij} unrestricted.

Alternatively, if the ordered responses are mirrored or symmetrically labeled, one can assume symmetric thresholds around zero such that the length of intervals for symmetrically labeled responses are the same. In this case, scale invariance can be achieved by fixing the length of one interval to an arbitrary number.

In this paper we fix the intercept terms $(\beta_{0j})_{j\in\{1,\dots,J\}}$ to zero and the variance of the errors to unity, such that $\boldsymbol{\Sigma}_i = \boldsymbol{R}_i$ becomes a correlation matrix. Moreover, in the parametric model we assume a sector specific correlation structure for the errors $\boldsymbol{R}_{g(i)}$, where $g(i)$ denotes the business sector of firm-year i. In other words, the correlation structure does not vary across subjects within the same business sector. In the presence of missing observations, $\boldsymbol{R}_{i,g(i)}$ the sub-matrix of $\boldsymbol{R}_{g(i)}$ denotes the correlation matrix corresponding to the underlying variables generating the observed outcomes $\boldsymbol{Y}_i = [Y_{ij}]_{j\in J_i}$ and is obtained by choosing the elements of $\boldsymbol{R}_{g(i)}$ corresponding to the available ratings (i.e., which lie in rows J_i and columns J_i).

3 Estimation

Let $\boldsymbol{\Gamma}$ denote the vector containing the threshold parameters, the regression coefficients, and the elements of the matrices $\boldsymbol{R}_{g(i)}$ to be estimated. The likelihood of the model is then given by the product:

$$\mathscr{L}(\boldsymbol{\Gamma}|\widetilde{X}, Y) = \prod_{i=1}^{n} \mathbb{P}(\cap_{j\in J_i} Y_{ij} = r_{ij}|\boldsymbol{\Gamma}, \widetilde{X})^{w_i} = \prod_{i=1}^{n} \left(\int_{D_i} f_{q_i}(\widetilde{\boldsymbol{Y}}_i|\boldsymbol{\Gamma}, \widetilde{X}) d^{q_i}\widetilde{\boldsymbol{Y}}_i \right)^{w_i},$$

where $D_i = \prod_{j\in J_i}(\theta_{j,r_{ij}-1}, \theta_{j,r_{ij}}]$ is a Cartesian product, w_i are non-negative weights, which in the simplest case are set to one, and f_{i,q_i} is the q_i-dimensional density corresponding to the distribution function F_{i,q_i}.

In order to estimate the model parameters we use a composite likelihood approach, where the full likelihood is approximated by a pseudo-likelihood which will be constructed from lower dimensional marginal distributions, more specifically by "aggregating" the likelihoods corresponding to pairs and triplets of observations, respectively. In the presence of observations missing completely at random, the composite likelihood will be constructed from the available outcomes for each firm-year i. If the number of outcomes q_i is less than two for the pairwise approach or three for the tripletwise approach, the marginal q_i-dimensional marginal probabilities are used. For the sake of notation we introduce an $n \times J$ binary index matrix Z, where each element z_{ij} takes a value of 1 if $j \in J_i$ and 0 otherwise. The pairwise log-likelihood is given by:

$$c\ell(\mathbf{\Gamma}|\widetilde{X}, Y) = \sum_{i=1}^{n} w_i \Bigg[\sum_{k=1}^{J-1}\sum_{l=k+1}^{J} \mathbb{1}_{\{z_{ik}z_{il}=1\}} \log\Big(\mathbb{P}(Y_{ik}=r_{ik}, Y_{il}=r_{il}|\mathbf{\Gamma}, \widetilde{X})\Big) + \\ \mathbb{1}_{\{q_i=1\}}\sum_{k=1}^{J} \mathbb{1}_{\{z_{ik}=1\}} \log\Big(\mathbb{P}(Y_{ik}=r_{ik}, |\mathbf{\Gamma}, \widetilde{X})\Big)\Bigg]. \tag{4}$$

Similarly, the tripletwise log-likelihood is:

$$c\ell(\mathbf{\Gamma}|\widetilde{X}, Y) = \sum_{i=1}^{n} w_i \Bigg[\sum_{k=1}^{J-2}\sum_{l=k+1}^{J-1}\sum_{m=l+1}^{J} \\ \mathbb{1}_{\{z_{ik}z_{il}z_{im}=1\}} \log\Big(\mathbb{P}(Y_{ik}=r_{ik}, Y_{il}=r_{il}, Y_{im}=r_{im}|\mathbf{\Gamma}, \widetilde{X})\Big) + \\ \mathbb{1}_{\{q_i=2\}}\sum_{k=1}^{J-1}\sum_{l=k+1}^{J} \mathbb{1}_{\{z_{ik}z_{il}=1\}} \log\Big(\mathbb{P}(Y_{ik}=r_{ik}, Y_{il}=r_{il}|\mathbf{\Gamma}, \widetilde{X})\Big) + \\ \mathbb{1}_{\{q_i=1\}}\sum_{k=1}^{J} \mathbb{1}_{\{z_{ik}=1\}} \log\Big(\mathbb{P}(Y_{ik}=r_{ik}, |\mathbf{\Gamma}, \widetilde{X})\Big)\Bigg]. \tag{5}$$

If, for the case of no missing observations, the errors follow a J-dimensional multivariate normal or multivariate logistic distribution, the lower dimensional marginal distributions F_{i,q_i} are also normally or logistically distributed. In the sequel we denote by $f_{i,1}$, $f_{i,2}$ and $f_{i,3}$ the uni-, bi- and trivariate densities corresponding to $F_{i,1}$, $F_{i,2}$ and $F_{i,3}$. Hence, the marginal probabilities can be expressed as:

$$\mathbb{P}(Y_{ik}=r_{ik}, Y_{il}=r_{il}, Y_{im}=r_{im}|\cdot) = \int_{L_{ik}}^{U_{ik}}\int_{L_{il}}^{U_{il}}\int_{L_{im}}^{U_{im}} f_{i,3}(v_{ik}, v_{il}, v_{im}|\cdot)dv_{ik}dv_{il}dv_{im},$$

$$\mathbb{P}(Y_{ik}=r_{ik}, Y_{il}=r_{il}|\cdot) = \int_{L_{ik}}^{U_{ik}}\int_{L_{il}}^{U_{il}} f_{i,2}(v_{ik}, v_{il}|\cdot)dv_{ik}dv_{il},$$

$$\mathbb{P}(Y_{ik}=r_{ik}|\cdot) = \int_{L_{ik}}^{U_{ik}} f_{i,1}(v_{ik})dv_{ik},$$

where $U_{ij} = \theta_{j,r_{ij}} - \widetilde{\boldsymbol{x}}_i^\top \widetilde{\boldsymbol{\beta}}_j$ denote the upper and $L_{ij} = \theta_{j,r_{ij}-1} - \widetilde{\boldsymbol{x}}_i^\top \widetilde{\boldsymbol{\beta}}_j$ the lower integration bounds.

Point maximum composite likelihood estimates $\hat{\mathbf{\Gamma}}_{\mathrm{CL}}$ are obtained by direct maximization using general purpose optimizers. In order to quantify the uncertainty of the maximum composite likelihood estimates, numerical differentiation techniques are used to compute the standard errors. Under certain regularity conditions, the maximum composite likelihood estimator is consistent as $n \to \infty$

and J fixed and asymptotically normal with asymptotic mean $\mathbf{\Gamma}$ and covariance matrix:

$$G(\mathbf{\Gamma})^{-1} = H^{-1}(\mathbf{\Gamma})V(\mathbf{\Gamma})H^{-1}(\mathbf{\Gamma}),$$

where $G(\mathbf{\Gamma})$ denotes the Godambe information matrix, $H(\mathbf{\Gamma})$ is the Hessian (sensitivity matrix) and $V(\mathbf{\Gamma})$ is the variability matrix (Varin, 2008). The sample estimates of $H(\mathbf{\Gamma})$ and $V(\mathbf{\Gamma})$ are given by:

$$\hat{V}(\mathbf{\Gamma}) = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial c\ell_i(\hat{\mathbf{\Gamma}}_{\mathrm{CL}}|\mathbf{Y}_i)}{\partial\mathbf{\Gamma}}\left(\frac{\partial c\ell_i(\hat{\mathbf{\Gamma}}_{\mathrm{CL}}|\mathbf{Y}_i)}{\partial\mathbf{\Gamma}}\right)^{\top}, \quad \hat{H}(\mathbf{\Gamma}) = -\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2 c\ell_i(\hat{\mathbf{\Gamma}}_{\mathrm{CL}}|\mathbf{Y}_i)}{\partial\mathbf{\Gamma}\partial\mathbf{\Gamma}^{\top}},$$

where $c\ell_i(\mathbf{\Gamma}|\mathbf{Y}_i)$ denotes the i-th component of the composite log-likelihood. For model comparison the composite likelihood information criterion can be used: $\mathrm{CLIC}(\mathbf{\Gamma}) = -2\ c\ell(\hat{\mathbf{\Gamma}}_{\mathrm{CL}}|X,Y) + k\ \mathrm{tr}(\hat{V}(\mathbf{\Gamma})\hat{H}(\mathbf{\Gamma})^{-1})$ (where $k = 2$ corresponds to CLIC-AIC and $k = \log(n)$ corresponds to CLIC-BIC).

To achieve monotonicity in the threshold parameters $\boldsymbol{\theta}_{j,\cdot}$ we set $\theta_{j,1} = \gamma_{j,1}$ and $\theta_{j,r} = \theta_{j,r-1} + \exp(\gamma_{j,r})$ for $r = 2,\ldots,K_j - 1$, and estimate the vector of unconstrained parameters $[\boldsymbol{\gamma}_{j,\cdot}]_{j\in\{1,\ldots,J\}}$. For all correlation matrices we use the spherical parametrization described in Pinheiro and Bates (1996) and transform the constrained parameter space into an unconstrained one. The spherical parametrization for covariance matrices has the advantage over other parametrizations in that it can easily be modified to apply to a correlation matrix.

4 Simulation Study

The aim of the simulation study is to investigate the following aspects: First, in order to assess how the sample size n influences the accuracy of the pairwise likelihood estimates, we simulate data sets with different numbers of observations and plot the mean squared errors of the estimates. Second, we investigate how the bias and the variance of the composite likelihood estimates changes when using the pairwise versus the tripletwise likelihood for both the probit and the logit links. Finally, motivated by the high number of incomplete outcomes in the credit ratings data set, we explore the performance of the pairwise likelihood in the presence of observations missing completely at random for three and five outcome variables. In addition, we include six groups of observations with different correlation patterns, which in the application case would correspond to business sectors.

For the probit link we simulate the error terms from the multivariate normal distribution. In order to simulate errors from the multivariate logistic distribution defined in Equation (3), we generate samples uniformly distributed on the $[0,1]$ interval from the t copula $(u_{i1},\ldots u_{iq_i})$. In order to obtain marginally logistic distributed errors ϵ_{ij} we use the transformation $\epsilon_{ij} = L^{-1}(u_{ij})$ where $L^{-1}(x)$ denotes the inverse univariate logistic distribution function.

In all settings, we work with three covariates for each outcome, which we simulate from a standard normal distribution and assume the vector of coefficients $\boldsymbol{\beta}_j = (1.2,-0.2,-1)^{\top}$ for all $j \in J$ outcomes. In our simulation study with $J = 3$ outcome variables, we use the following set of threshold parameters: three thresholds for the first outcome $\boldsymbol{\theta}_1 = (-1,0,1)^{\top}$, three thresholds for outcome two $\boldsymbol{\theta}_2 = (-2,0,2)^{\top}$ and five thresholds for the third outcome $\boldsymbol{\theta}_3 = (-1.5,-0.5,0,0.5,1.5)^{\top}$. The underlying error terms are assumed to have different degrees of correlation. More details are provided for each simulation exercise in the following subsections.

In the simulation study, we follow Bhat et al. (2010) and proceed in the following way:

1. Simulate S data sets with n subjects, where each subject i has J outcome variables.
2. Estimate the composite likelihood parameters for each data set and compute the mean estimate for all parameters.
3. Estimate the asymptotic standard errors using the Godambe information matrix for each data set and compute the mean[3] for all parameters.
4. Compute bias and absolute percentage bias (APB)[4]:
$$\text{APB} = \left| \frac{\text{true parameter} \; - \; \text{mean estimate}}{\text{true parameter}} \right|.$$
5. Compute the finite sample error through calculating the standard deviation across all S data sets for each parameter.
6. Calculate a relative efficiency measure of estimator 2 compared to estimator 1
$$\text{RE} = \frac{\text{se}_1}{\text{se}_2}.$$
for both, the asymptotic as well as the finite sample standard errors.

4.1 Investigating the effect of the sample size on the pairwise likelihood estimates

In this part we investigate the influence of the number of subjects n on the pairwise likelihood estimates for both the probit and the logit link. For this purpose, we use three different correlation structures and simulate for each correlation pattern $S = 100$ data sets for increasing number of subjects n. We use a high correlation ($\boldsymbol{R}_1$; solid line), a moderate correlation ($\boldsymbol{R}_2$; dashed line) and a low correlation matrix ($\boldsymbol{R}_3$; dotted line). The correlation matrices can be found in Subsection 4.3. In Figure 1 average mean squared errors (MSE) are plotted against the number of subjects n. The average MSEs of the threshold parameters as well as the coefficients show, as expected, no difference between the data sets simulated with different correlation structures. Conversely, the MSEs of the correlation parameters differ across different degrees of correlation. We observe that correlation parameters of the high correlation data sets are recovered better compared to the moderate and low correlation ones. This finding has been previously reported also by e.g., Bhat et al. (2010) in their simulation study for the multivariate probit model. The last plot shows the average MSEs of all estimated parameters indicating that from $n = 500$ subjects the MSE curves start to flatten out. MSEs are in general low and even for smaller sample sizes (like $n = 100$) we obtain reasonable results. On average the logit link MSEs are slightly higher than the ones obtained by probit link, but this seems to not be the case for the correlation parameters.

[3]With one exception: In the case of the tripletwise estimates we compute the median due to instabilities in the numerical derivatives of the trivariate normal distribution function.

[4]If the true parameter is zero we do not report the APB.

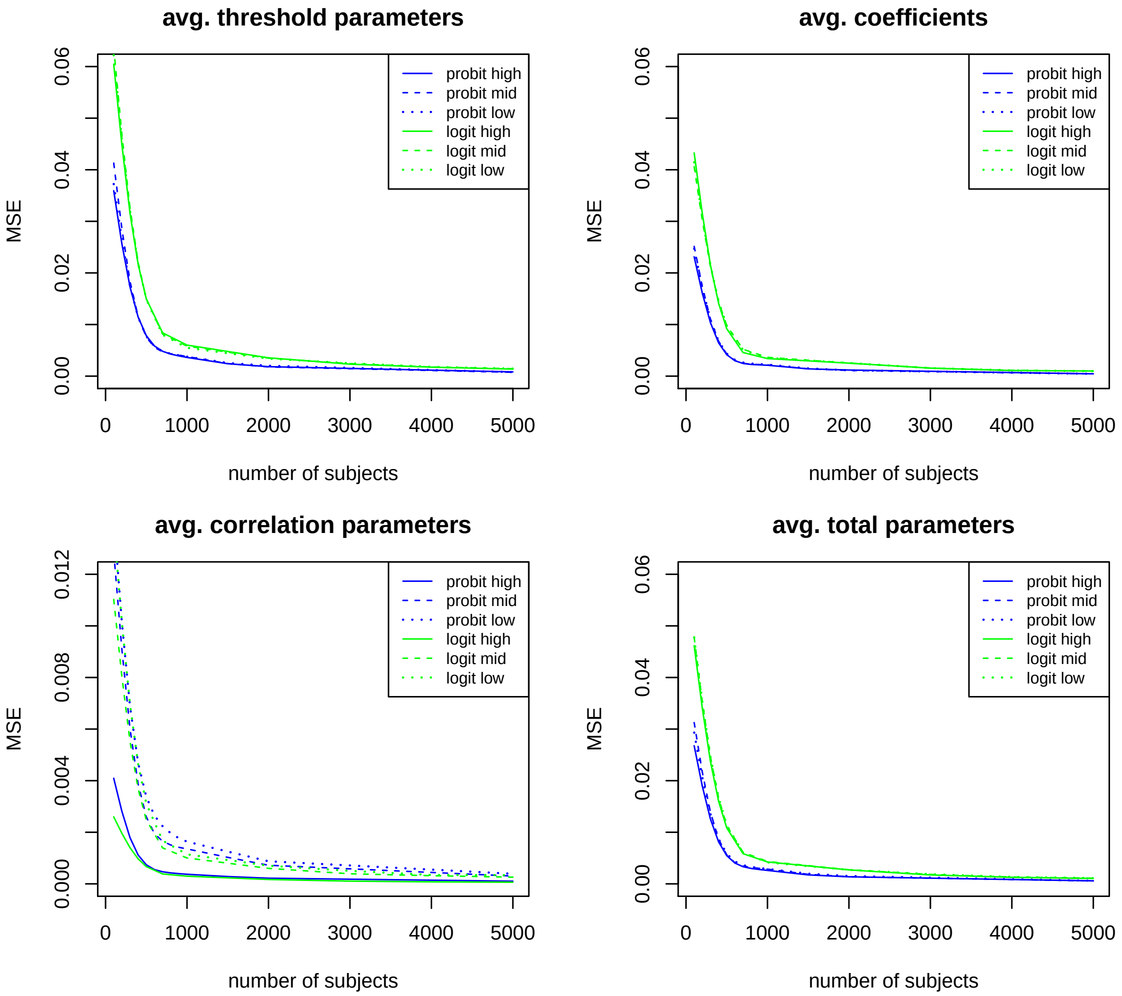

Figure 1: These plots display the averaged MSEs for increasing number of subjects n for the **probit link** (blue) and the **logit link** (green).

We decide to proceed in the sequel of the paper with $n = 1000$ subjects per group, also motivated by the application case where the smallest business sector contains around 1000 subjects.

4.2 Comparison pairwise vs. tripletwise likelihood

In order to compare pairwise and tripletwise likelihood estimators we simulate $S = 1000$ data sets with $n = 1000$ subjects and three outcome variables ($J = 3$). Table B.1 (probit link) and Table B.2 (logit link) present a comparison between the pairwise and tripletwise likelihood (which for $J = 3$ represents the full likelihood) estimates. For each link, both approaches seem to recover all parameters very well. For the probit link, comparing the APB of the two estimation approaches yields a range from 0.05% to 0.93% for the pairwise and a range from 0.00% to 0.89% for the tripletwise (or full) likelihood approach. In this case, the relative efficiency of the tripletwise estimators to the pairwise estimators is close to one for asymptotic as well as finite sample standard errors.

For the logit link the APB ranges from 0.04% to 2.15% for the pairwise approach and from 0.02% to 2.08% for the tripletwise approach. The relative efficiency measure is again close to one. For both link functions the asymptotic standard errors are close to the finite sample standard errors. For the logit link the standard errors of the threshold and coefficient parameters are higher than for the probit link, while for the correlation parameters this difference disappears. An inspection of the QQ-plots for the pairwise and tripletwise parameter estimates reveals that the empirical distribution of the $S = 1000$ estimates is well approximated by a normal distribution.

According to the results, there seems to be no substantial improvement in the parameter estimates when using the tripletwise approach. In terms of computing time, the pairwise likelihood approach (on average 263.68 seconds per data set) outperforms the tripletwise likelihood approach (on average 935.54 seconds per data set) by a factor of 3.5. Computations have been performed on 25 IBM dx360M3 nodes within a cluster of workstations.

In addition, when moving from two to three dimensions estimating the asymptotic standard errors can be problematic. This is because the numerical computation of the gradient and Hessian of the objective function highly depends on the algorithm used for computing the multivariate normal or t probabilities, which again delivers an approximation and must rely on deterministic methods (otherwise the derivatives do not exist). According to our simulations, for two dimensions the procedure is stable, but in more than two dimensions it can lead to numerical instabilities. Given the similar performance, computing time and stability of the numerical estimation of the standard errors, we decide to use the pairwise likelihood approach for the analysis of the multiple credit ratings data set in Section 5.

4.3 Simulation study with three outcomes and six different sector correlations

In this subsection we analyze the performance of the pairwise likelihood approach in the presence of missing observations for three outcome variables. We simulate $S = 1000$ data sets with $n = 6000$ subjects, where each subject i has three outcome variables ($J = 3$) yielding in total 18000 observations in the outcome variables. We allow for 6 different sectors with each $n_s = 1000$ subjects and choose two high correlation, two moderate correlation and two low correlation matrices:

$$\boldsymbol{R}_1 = \begin{pmatrix} 1.0 & 0.8 & 0.7 \\ 0.8 & 1.0 & 0.9 \\ 0.7 & 0.9 & 1.0 \end{pmatrix}, \qquad \boldsymbol{R}_2 = \begin{pmatrix} 1.0 & 0.5 & 0.3 \\ 0.5 & 1.0 & 0.4 \\ 0.3 & 0.4 & 1.0 \end{pmatrix}, \qquad \boldsymbol{R}_3 = \begin{pmatrix} 1.0 & 0.2 & 0.3 \\ 0.2 & 1.0 & 0.1 \\ 0.3 & 0.1 & 1.0 \end{pmatrix},$$

$$\boldsymbol{R}_4 = \begin{pmatrix} 1.0 & 0.9 & 0.9 \\ 0.9 & 1.0 & 0.9 \\ 0.9 & 0.9 & 1.0 \end{pmatrix}, \qquad \boldsymbol{R}_5 = \begin{pmatrix} 1.0 & 0.8 & 0.3 \\ 0.8 & 1.0 & 0.6 \\ 0.3 & 0.6 & 1.0 \end{pmatrix}, \qquad \boldsymbol{R}_6 = \begin{pmatrix} 1.0 & 0.1 & 0.1 \\ 0.1 & 1.0 & 0.1 \\ 0.1 & 0.1 & 1.0 \end{pmatrix}.$$

For the probit link, Table B.3 presents the parameter estimates of both the full observations model and the model containing missing observations. The results for the logit link are displayed in the Table B.4.

Full observations model In the full observations model we observe excellent estimates for all parameters. In particular for the probit link, the threshold parameters and coefficients are recovered very well. The APB ranges from 0.04% to 0.30%. In the case of correlation parameters we observe that high correlation parameters are recovered extremely well (APB between 0.01% and 0.16%), in contrast to low correlation parameters, where we observe higher APB (up to 1.90%) as well as standard errors. Even though the model performs better for high correlation structures, we can conclude that pairwise likelihood estimates are very good for different correlation patterns. In the presence of the logit link we observe slightly worse regression coefficient (APB from 0.58% to 2.53%) and threshold estimates (APB from 0.32% to 0.67%), but slightly better estimates for low and moderate correlations (APB from 0.01% to 1.13%) compared to the probit link.

Missing observations model We repeated the simulation this time with observations missing completely at random in the outcome variables of the simulated data sets. We randomly remove 5% of the first outcome variable, 20% of the second outcome and 50% of the third outcome. Overall for both link functions, all parameter estimates are recovered very well in the missing observation model. In analogy to the full observations model with probit link, the threshold and coefficient parameters have an APB ranging from 0.03% to 0.36%. High correlation parameters (APB from 0.03% to 0.19%) are recovered better compared to low correlation parameters (APB up to 2.83%). In addition, standard errors increase for all parameters with the number of missing observations. In the logit model with missing observations, the threshold and coefficient parameters as well as the high correlation parameters are recovered very well, in contrast to low correlation parameters, where we observe that missing observations have an impact on the quality of the estimates (ABP increases up to 6.35%).

Full observations model vs. Missing observations model First, we compare the parameter estimates of the full and the missing observations model with probit link. As expected, we observe smaller APB and standard errors for almost all parameters in the full model. In case of threshold parameters and coefficients, we do not observe a big difference in the pairwise likelihood estimates. While large correlation parameters are recovered very well in both models, we observe a significant impact of missing observations on the estimation quality of low correlation parameters (e.g. APB increases from 1.90% to 2.83% for parameter ρ_{23}^3).

Nevertheless, even if we omit 50% of the observations of one particular outcome variable, all parameter estimates remain very good as long as the number of remaining observations is not too low. In terms of relative efficiency our measure yields approximately 0.9 for most parameters corresponding to the outcome with 5% missing observations, approximately 0.85 for parameters corresponding to outcome two (20% missing observations) and approximately 0.7 for parameters corresponding to the third outcome with 50% of missing observations. Moreover, a comparison for the logit link models shows similar aspects. For threshold as well as coefficient estimates, the estimation quality does not suffer strongly in the presence of missing observations. The quality of the correlation parameters is only affected in dimensions with a lot of missings and low correlation (e.g., correlation parameters between the second and third outcome). In such a case, the maximal APB increases from 2.49% up to 6.35%. In summary, we are confident that, even though one has to deal with outcomes with high percentage of missing values, the pairwise likelihood estimates can still recover the parameters of interest in a reliable way.

Simulation study with five outcomes In addition, a simulation study with $J = 5$ outcomes is conducted. Again, pairwise likelihood estimates of a full observations model and a model with outcomes containing up to 70% missing observations are estimated. The findings are similar to the ones of the three dimensional simulation study and we observe that all parameters are recovered very well in models with a higher number of outcomes. More information can be found in Appendix A.

5 Multivariate Analysis of Credit Ratings

We base our empirical analysis on a data set of US firms rated by S&P, Moody's and Fitch over the period 2000–2013. We chose this time frame as Fitch became an established player in the US ratings market around year 2000 (Becker and Milbourn, 2011).

5.1 Data

We collect historical long-term issuer credit ratings from S&P, Moody's and Fitch, the three biggest CRAs in the US market. S&P domestic long-term issuer credit ratings are retrieved from the S&P Capital IQ's Compustat North America© Ratings file, while issuer credit ratings from Moody's and Fitch were provided by the CRAs themselves. The CRAs assign ratings on an ordinal scale. S&P and Fitch assign issuers to 21 non-default categories[5]. Moody's rating system for issuers comprises 20 non-default rating classes and uses different labeling[6], where AAA and Aaa, respectively represent the highest credit quality and hence lowest default risk. Firms falling into the best ten categories (AAA/Aaa to $BBB{-}/Baa3$) are considered investment grade (IG) firms, while those falling into $BB{+}/Ba1$ to C/Ca are speculative grade (SG) firms.

In order to build the covariates, annual financial statement data and daily stock prices from the Center of Research in Security Prices (CRSP) are downloaded for the S&P Capital IQ's Compustat North America© universe of publicly traded US companies. Following the existing literature (e.g., Shumway, 2001; Campbell et al., 2008; Alp, 2013) and the rating methodology published by the CRAs (Puccia et al., 2013; Tennant et al., 2007; Hunter et al., 2014), we build the following covariates: *free operating cash-flow coverage ratio ([operating cash-flow − capital expenditures + interest expenses]/interest expenses)*, *cash/assets*, *tangibility (fixed assets/assets)*, *debt/assets*, *short-term debt/debt*, *retained earnings/assets*, *return on capital (earnings before interest and taxes/equity and debt)*, *earnings before interest and taxes/sales*, *research and development expenses (R&D)/assets* and *capital expenditures/assets*. In addition, we use daily stock prices to compute the following measures: *relative size (RSIZE)* is the logarithm of the ratio of market value of equity (computed as the

[5] AAA, $AA+$, AA, $AA-$, $A+$, A, $A-$, $BBB+$, BBB, $BBB-$, $BB+$, BB, $BB-$, $B+$, B, $B-$, $CCC+$, CCC, $CCC-$, CC and C.

[6] Aaa, $Aa1$, $Aa2$, $Aa3$, $A1$, $A2$, $A3$, $Baa1$, $Baa2$, $Baa3$, $Ba1$, $Ba2$, $Ba3$, $B1$, $B2$, $B3$, $Caa1$, $Caa2$, $Caa3$, Ca

Table 1: This table displays in the diagonal the number of ratings from the three CRAs in our data set from 2000 to 2013. The off-diagonal displays the number of co-rated firms.

	S&P	Moody's	Fitch
S&P	19874	12270	4173
Moody's		13168	3749
Fitch			4365

average stock price in the year previous to the observation times the number of shares outstanding) to the average value of the CRSP value weighted index. *BETA* is a measure of systematic risk, which represents the relative volatility of a stock price compared to the overall market. *SIGMA* is a measure of idiosyncratic risk. We regress the daily stock price in the year before the observation on the daily CRSP value weighted index. *BETA* is the regression coefficient and *SIGMA* is the standard deviation of the residuals of this regression. The last measure is the *market assets to book assets ratio (MB)* which is market equity plus book liabilities divided by book assets.

We follow standard practice in the literature and remove financials (GICS code 40) and utilities (GICS code 55) from the sample, as these firms have a special regime of reporting their annual figures and this fact might distort the results. We match the ratings data with financial statement data from Compustat using CUSIPs. To ensure that these data are observable to the rating agencies at the time the rating is issued, we match each rating with financial statement data lagged by three months. We choose the three months lag, as all publicly traded US companies must file their annual reports with the Securities and Exchange Commission within 90 days of the fiscal year end.

The merged sample consists of 20880 firm-year observations and 2876 companies for which at least one rating is available. Table 1 shows the number of non-missing ratings and co-ratings between the CRAs. S&P rates 95%, Moody's 63% and Fitch only 21% of the firm-year observations in the sample. Only 3665 firm-years (18%) have a rating from all three CRAs. Figure 2 shows the distributions of the ratings for each CRA. For further analysis we aggregate the "+" and "−" ratings for S&P and Fitch and the "1" and "3" ratings for Moody's to the middle rating. Moreover, following the practice of the CRAs in their report series, we aggregate classes *CCC* to *C* for S&P and Fitch. The distribution of the ratings using the aggregated scale is presented in Figure 3.

We winsorize all variables at the 97.5% quantile and additionally the variables which can take negative values at the 2.5% quantile. Missing values in the ratios are replaced by the sectorwise median in each year. Given the different scale of the covariates we standardize them to have mean zero and variance equal to one.

In order to perform a sectorwise correlation analysis, firms are classified into business sectors according to the Global Industry Classification Standard (GICS). We use eight sectors in the analysis: energy (GICS code 10, 2617 observations), materials (GICS code 15, 2482 observations), industrials (GICS code 20, 3631 observations), consumer discretionary (GICS code 25, 5176 observations), consumer staples (GICS code 30, 1646 observations), health care (GICS code 35, 1955 observations), information technology (GICS code 45, 2194 observations) and telecommunication services (GICS

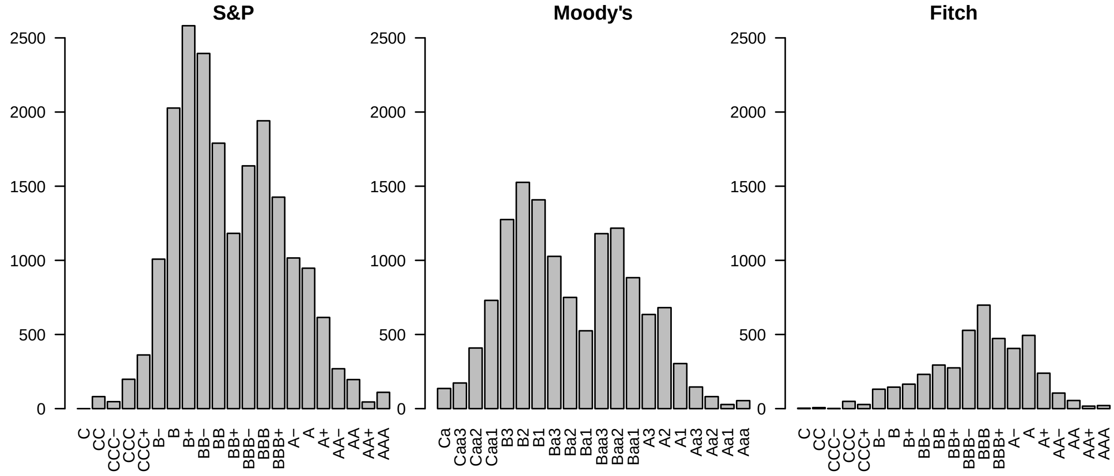

Figure 2: This figure displays the distribution of ratings in the original scale containing 21 rating classes.

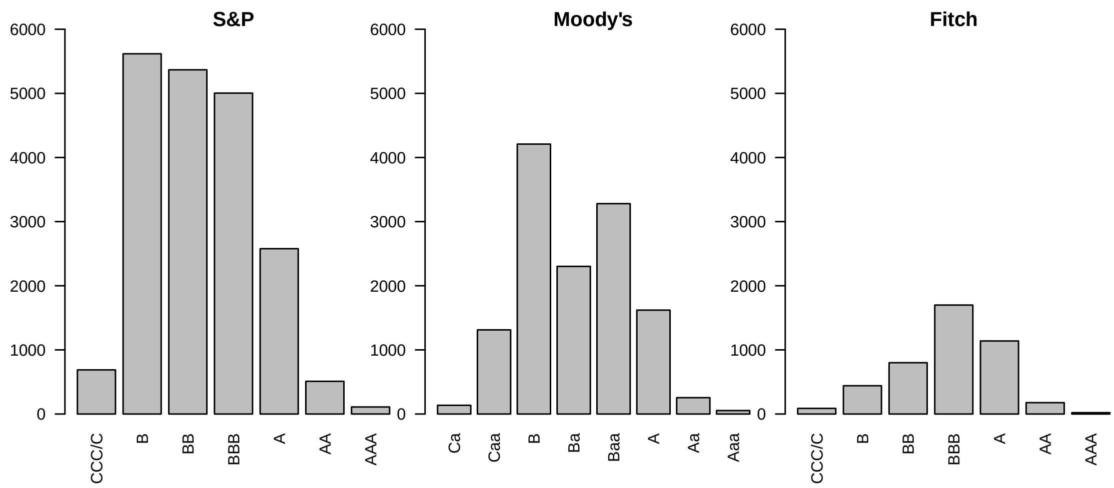

Figure 3: This figure displays the distribution of ratings in the aggregated scale containing 7 rating classes for S&P and Fitch and 8 rating classes for Moody's

code 50, 1179 observations).

5.2 Results

Model (1) is fitted to the ratings data set. The latent variable motivation in ordinal models is an intuitive setting for the application case. In the context of credit risk one may think of the underlying latent variable as the latent creditworthiness of a firm, which is measured on a continuous scale. In the literature, this latent variable has been introduced under different names and in different settings. For example, Altman (1968) introduced the Z-score, a linear combination of multiple accounting ratios, as a measure to predict corporate defaults. Furthermore, in his seminal work, Merton (1974) proxies creditworthiness by the distance-to-default, which measures the distance of the firm's log asset value to its default threshold on the real line. Ratings can then be considered as a coarser version of this latent variable. Low values of the latent creditworthiness will translate to the worst rating classes, while the right tail of the distribution of the latent variables will correspond to the best rating classes.

We use both the probit and the logit links in the estimation of the model. The CLIC-BIC for the model using the logit link is slightly lower (91920.3) than the CLIC-BIC for the probit model (94089.4). We therefore proceed in the following only with the discussion of the results using the logit link. The results of the multivariate probit model can be found Tables B.5 and B.6 in Appendix. No notable differences occur. As expected, the threshold and regression coefficients for the logit model are larger than the ones of the probit model by a factor of approximately $\sqrt{\pi^2/3}$.

It is to be noted that the estimated thresholds and coefficients represent signal to noise ratios due to identifiability constraints and one needs to proceed with care when interpreting the results. Nonetheless, the results provide several interesting insights.

Threshold parameters The estimated threshold parameters together with their standard errors for the multivariate logit model are presented in Table 2. Moody's seems to be the most conservative rater, with almost all threshold parameters higher than the other two CRAs. While for the investment grade classes the difference between S&P and Moody's thresholds is relatively small, this is not the case for the speculative grade rating classes, where Moody's seems to distance itself from S&P in the way it assigns ratings and tends to be more conservative. Fitch on the other hand has significantly lower threshold parameters $BBB|A$ and $BB|BBB$ than S&P, which could translate into a more optimistic rating scale around the investment–speculative grade frontier.

Regression coefficients Table 3 presents the regression coefficients. Firms with higher free operating cash-flow coverage ratios, more tangible assets, higher proportion of short-term debt (which is less risky than long-term debt), high profitability (measured by retained earnings to assets, return on capital or EBIT/sales), which spend more on R&D and have a bigger size tend to get better ratings. On the other hand, firms with higher debt ratios, capital expenditures, idiosyncratic and systematic risk tend to get worse credit ratings. Moreover, high liquidity levels are inversely related to creditworthiness which is rather counter intuitive. This is in line with previous results and can be explained by the fact that a conservative cash policy is more likely to be pursued by a firm that

Table 2: This table displays the threshold parameter estimates from the multivariate ordered **logit** model using the multiple corporate credit ratings data set.

	S&P		Fitch			Moody's	
Thresholds	Est.	SE	Est.	SE	Thresholds	Est.	SE
					Ca|Caa	−8.2069	0.1585
CCC/C|B	−6.7774	0.1004	−6.1065	0.1902	Caa|B	−4.9643	0.0882
B|BB	−2.6980	0.0655	−2.9145	0.1181	B|Ba	−1.8586	0.0677
BB|BBB	−0.5638	0.0632	−0.9771	0.1084	Ba|Baa	−0.4358	0.0671
BBB|A	1.8639	0.0666	1.4639	0.1096	Baa|A	2.0049	0.0710
A|AA	4.5183	0.0821	4.3426	0.1275	A|Aa	4.7073	0.0911
AA|AAA	6.6068	0.1228	6.8602	0.2050	Aa|Aaa	6.8622	0.1419

Table 3: This table displays the regression coefficients from the multivariate ordered **logit** model using the multiple corporate credit ratings data set.

	S&P		Moody's		Fitch	
Covariate	Est.	SE	Est.	SE	Est.	SE
operating CF cov.	0.0888	0.0224	0.1139	0.0242	0.0793	0.0317
cash/assets	−0.1170	0.0176	−0.1025	0.0188	−0.1463	0.0245
tangibility	0.2674	0.0207	0.3043	0.0228	0.2229	0.0288
debt/assets	−0.6988	0.0236	−0.6495	0.0258	−0.8313	0.0375
ST debt/debt	0.2085	0.0199	0.2440	0.0225	0.2509	0.0296
ret.earnings/assets	0.7554	0.0232	0.7402	0.0262	0.6711	0.0317
return on capital	0.3403	0.0228	0.3592	0.0246	0.3591	0.0317
EBIT/sales	0.2146	0.0202	0.1975	0.0206	0.2167	0.0258
R&D/assets	0.2737	0.0177	0.2544	0.0191	0.2790	0.0242
capex/assets	−0.1325	0.0208	−0.1715	0.0226	−0.0840	0.0328
RSIZE	0.9399	0.0221	1.0191	0.0246	0.8234	0.0304
BETA	−0.2069	0.0178	−0.1772	0.0189	−0.2165	0.0252
SIGMA	−0.6476	0.0288	−0.6086	0.0311	−0.5940	0.0458
MB	−0.2275	0.0203	−0.1997	0.0233	−0.1130	0.0299

finds itself close to distress and that higher cash holdings increase the long-term probability of default (e.g., Acharya et al., 2012; Alp, 2013). The market-to-book ratio (MB) is also inversely related to creditworthiness. This has also been found by Campbell et al. (2008), who argue that high MB ratio can point towards overvaluation of the firm in the market, which in turn can be a bad sign in terms of credit quality.

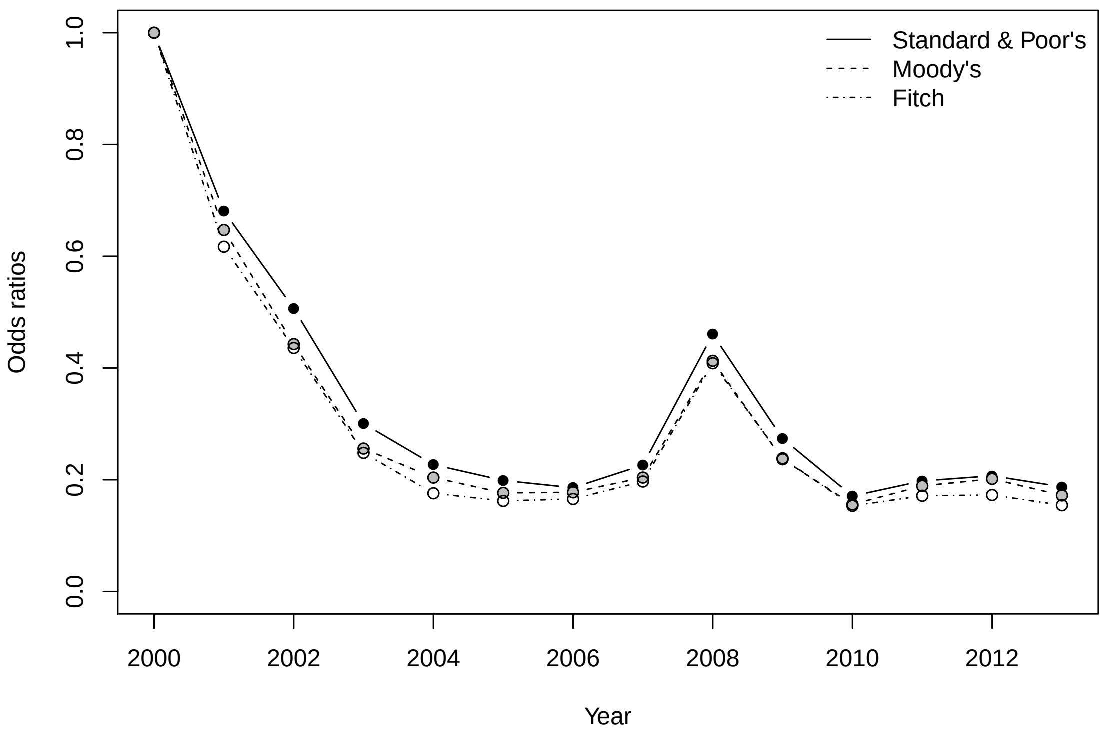

Figure 4: This figure displays the time dummy coefficients from 2000 to 2013 from the multivariate ordered **logit** model.

Year intercepts As previously mentioned, using the logit link has the advantage that the regression coefficients can be interpreted as marginal log odds ratios. For the year intercepts, this means that, for each year t and rater j, the odds of $Y \geq r$ against $Y < r$ (i.e., the odds of a firm being assigned to rating class r or better rather than in a worse class than r, for all r) are $\exp(\alpha_{tj})$ times the odds in 2000 (which is the baseline year), ceteris paribus. Figure 4 shows these odds ratios corresponding to the coefficients of the year dummies for each rating agency. We observe that the odds ratios are less than one for all years, which means that the odds of a firm with constant characteristics to get a better rating decrease after 2000. This can indicate a tightening of rating standards (also found by Alp, 2013). An interesting remark is that before the financial crisis the odds start increasing, reaching a peak in 2008 for all CRAs, indicating a possible loosening of the rating standards in the financial crisis. After 2008, the odds return and stabilize close to the levels before the financial crisis. In addition, the odds of Moody's and Fitch are nearly identical during the crisis.

Correlation parameters Figure 5 shows the estimated correlation parameters together with their standard errors. We interpret the correlations as measures of association between the three CRAs, even though they are often interpreted as measures of agreement. In general, we observe very high levels of association for all business sectors. In particular, very high levels of association for all three CRAs are identified for sectors like energy, materials, industrials and consumer staples. Other sectors like consumer discretionary, health care, information technology or telecommunication show small

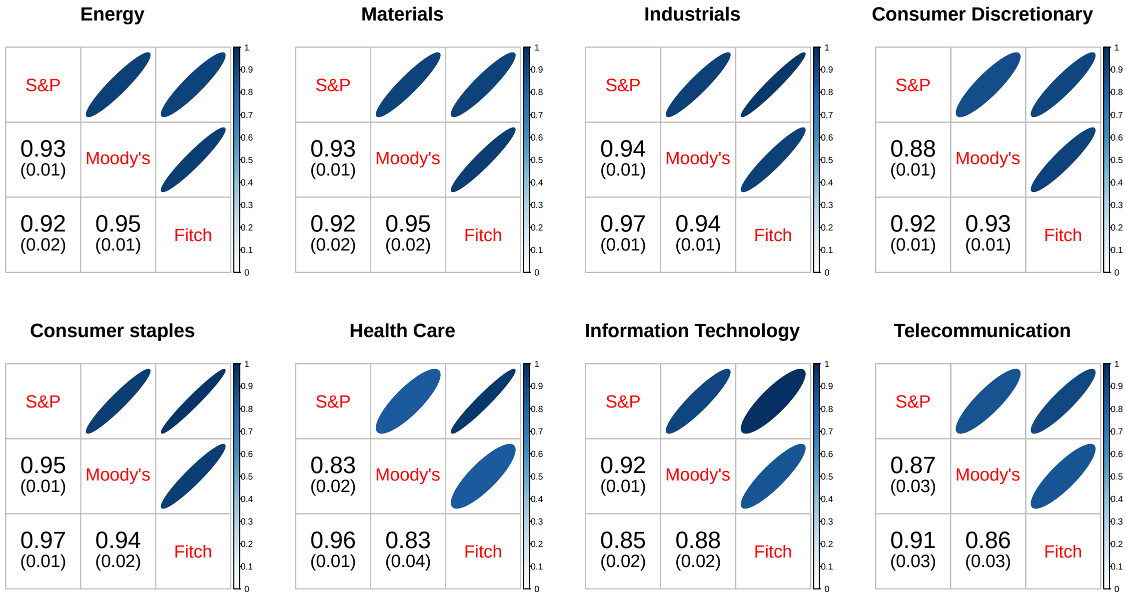

Figure 5: This figure shows the correlation estimates from the multivariate ordered **logit** model for different business sectors using the multiple corporate credit ratings data set. The standard errors are given in parentheses.

deviations in the association levels among the CRAs and exhibit correlations under 0.9. The high degree of correlation is good news, as it implies that firms have little incentives to engage in ratings shopping. Ratings "shopping" emerges when CRAs do not perfectly agree on the credit quality of a firm, as firms could exploit the disagreement by "shopping" the most favorable ratings (see for example Cantor and Packer, 1997; Becker and Milbourn, 2011; Bongaerts et al., 2012).

6 Concluding Remarks

In this paper we consider multivariate ordinal regression models with a latent variable specification in a credit risk context. This joint modeling approach is motivated by the case where multiple CRAs assess a firm's credit quality based on firm-level and market information and assign ordinal credit ratings accordingly. Composite likelihood methods are applied to estimate the model parameters and a simulation study is performed in order to investigate several aspects. First, we check how the sample size affects the pairwise likelihood estimates. We find that results are reasonable already for small sample sizes (e.g., 100 subjects) and that the MSEs flatten out for samples sizes higher than 500. For both link functions, high correlation parameters are better recovered than low correlation parameters, even though it seems that the logit link does a slightly better job at recovering low correlations. Second, we find that for three ordinal outcomes, using the pairwise approach has advantages over the tripletwise (or full) likelihood approach. Even though the tripletwise approach delivers slightly better estimates in terms of bias, the differences between the estimates are minimal and the pairwise approach is significantly faster than the tripletwise approach. Another relevant aspect for the application case, where the panel of credit ratings has many missing values especially for Fitch, is the influence of ignorable missing values on the pairwise likelihood estimates. We find that these estimates are robust to observations missing completely at random and threshold parameters, coefficients and high correlation parameters are all recovered very well. Low correlation dimensions are more sensitive to missing observations but, as long as the sample size is not too small, estimates are reliable. Additionally, a simulation study with five outcome variables was performed and similar results as for the three-dimensional case were observed. Simulation results are satisfactory for both the probit and the logit link functions.

In the empirical application, corporate credit ratings from S&P, Moody's and Fitch are matched to financial statement and stock price data for US publicly traded firms between 2000 and 2013. Relevant covariates which have an impact on the creditworthiness of firms are chosen according to prior literature. Moreover, we include time dummies in the analysis to capture changes in the rating standards over time. Association between the ordinal credit ratings is reflected in the correlation between the latent creditworthiness processes, which in our model depends on the business sector of the firm. We allow for different threshold parameters for each CRA and observe that Moody's tends to have a more conservative behavior, especially in the speculative grade classes, while Fitch seems to assign on average better ratings around the investment–speculative grade frontier. Moreover, all covariates have the expected sign and are consistent with the existing literature. We conclude that firms with higher debt ratio, capital expenditures, idiosyncratic and systematic risk, market to book ratio tend to get worse credit ratings. Larger, more profitable firms, which spend more on R&D and have high cash-flow coverage ratios, a higher proportion of tangible assets and of short-term debt tend to obtain better ratings. The coefficients of the year dummies indicate that rating standards in the sample period became stricter relative to the standards in 2000. This continuous "tightening" trend after 2000 was interrupted by a "loosening" of the standards during the financial crisis 2007–2009, but after 2010 the coefficients returned to the level before the crisis. The degree of inter-rater association for all business sectors is very high. Marginal differences are observed for few business sectors.

Possible extensions of this work include the incorporation of multi-level dependencies, such as time dependencies in the error terms and/or the implementation of different covariates in the error correlation matrix. The empirical analysis could be extended to incorporate additional ratings from smaller players in the US ratings market.

Computational details

All computations have been performed in R (R Core Team, 2017). For the computation of the bi- and trivariate normal and t probabilities we used the R package **mnormt** (Azzalini and Genz, 2016). The minimization of the negative log-likelihood has been performed by using the general purpose optimizers implemented in the package **optimx** (Nash and Varadhan, 2011; Nash, 2014). After trying all available solvers, we chose the NEWUOA solver (Powell, 2006), as it exhibited the highest convergence speed and also converged in all the simulation exercises. The numerical derivatives have been computed with the R package **numDeriv** (Gilbert and Varadhan, 2015).

References

Viral Acharya, Sergei A. Davydenko, and Ilya A. Strebulaev. Cash holdings and credit risk. *Review of Financial Studies*, 25(12):3572–3604, 2012.

Alan Agresti. *Analysis of ordinal categorical data.* Wiley Series in Probability and Statistics. John Wiley & Sons, 2 edition, 2010.

James H. Albert and Siddhartha Chib. Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88(422):669–679, 1993.

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

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

Adelchi Azzalini and Alan Genz. *The R package `mnormt`: The multivariate normal and t distributions (version 1.5-4)*, 2016.

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.

Bo Becker and Todd Milbourn. How did increased competition affect credit ratings? *Journal of Financial Economics*, 101(3):493–514, 2011.

Chandra R. Bhat and Sivaramakrishnan Srinivasan. A multidimensional mixed ordered-response model for analyzing weekend activity participation. *Transportation Research Part B: Methodological*, 39(3):255–278, 2005.

Chandra R. Bhat, Cristiano Varin, and Nazneen Ferdous. A comparison of the maximum simulated likelihood and composite marginal likelihood estimation approaches in the context of the multivariate ordered-response model. In William Greene and R. Carter Hill, editors, *Maximum Simulated Likelihood Methods and Applications*, volume 26 of *Advances in Econometrics*, pages 65–106. Emerald Group Publishing Limited, 2010.

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.

Dion Bongaerts, K. J. Martin Cremers, and William N. Goetzmann. Tiebreaker: Certification and multiple credit ratings. *The Journal of Finance*, 67(1):113–152, 2012.

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

Richard Cantor and Frank Packer. Differences of opinion and selection bias in the credit rating industry. *Journal of Banking & Finance*, 21(10):1395–1417, 1997.

Ming-Hui Chen and Dipak K. Dey. Bayesian analysis for correlated ordinal data models. In *Generalized linear models: A Bayesian perspective*, pages 135–162. New York: Marcel Dekker, 2000.

Siddhartha Chib and Edward Greenberg. Analysis of multivariate probit models. *Biometrika*, 85(2): 347–361, 1998.

Maria DeYoreo and Athanasios Kottas. Bayesian nonparametric modeling for multivariate ordinal regression. *ArXiv e-prints*, August 2014.

Paul Gilbert and Ravi Varadhan. *numDeriv: Accurate Numerical Derivatives*, 2015. R package version 2014.2-1.

William H. Greene and David A. Hensher. *Modeling Ordered Choices: A Primer*. Cambridge University Press, 2010.

Denitsa Grigorova, Elitsa Encheva, and Ralitza Gueorguieva. EM Algorithm for MLE of a Probit Model for Multiple Ordinal Outcomes. *Serdica Journal of Computing*, 7(3):227–244, 2013.

Emil J. Gumbel. Bivariate logistic distributions. *Journal of the American Statistical Association*, 56 (294):335–349, 1961.

Richard Hunter, Michael Dunning, Michael Simonton, Daniel Kastholm, and Andrew Steel. Corporate rating methodology. including short-term ratings and parent and subsidiary linkage. Technical report, Fitch Ratings, May 2014.

Hiroyuki Kawakatsu and Ann G. Largey. EM algorithms for ordered probit models with endogenous regressors. *The Econometrics Journal*, 12(1):164–186, 2009.

Jung Seek Kim and Brian T. Ratchford. A Bayesian multivariate probit for ordinal data with semiparametric random-effects. *Computational Statistics & Data Analysis*, 64:192–208, 2013.

Henrick J. Malik and Bovas Abraham. Multivariate logistic distributions. *The Annals of Statistics*, 1(3):588–590, 1973.

Robert C. Merton. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, 29(2):449–470, 1974.

John C. Nash. On best practice optimization methods in R. *Journal of Statistical Software*, 60(2): 1–14, 2014.

John C. Nash and Ravi Varadhan. Unifying optimization algorithms to aid software system users: optimx for R. *Journal of Statistical Software*, 43(9):1–14, 2011.

Nazanin Nooraee, Fentaw Abegaz, Johan Ormel, Ernst Wit, and Edwin R. van den Heuvel. An approximate marginal logistic distribution for the analysis of longitudinal ordinal data. *Biometrics*, 72(1):253–261, 2016.

Sean M. O'Brien and David B. Dunson. Bayesian multivariate logistic regression. *Biometrics*, 60(3): 739–746, 2004.

Kenne Pagui, Euloge Clovis, and Antonio Canale. Pairwise likelihood inference for multivariate ordinal responses with applications to customer satisfaction. *Applied Stochastic Models in Business and Industry*, 2015.

José C. Pinheiro and Douglas M. Bates. Unconstrained parametrizations for variance-covariance matrices. *Statistics and Computing*, 6(3):289–296, 1996.

Michael J. D. Powell. The NEWUOA software for unconstrained optimization without derivatives. In *Large-scale nonlinear optimization*, pages 255–297. Springer, 2006.

Mark Puccia, Lucy A. Collett, Peter Kernan, Andrew D. Palmer, Mark S. Mettrick, and Guy Deslondes. Request for comment: Corporate criteria. Technical report, Standard and Poor's Rating Services, June 2013.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL https://www.R-project.org/.

Darren M. Scott and Pavlos S. Kanaroglou. An activity-episode generation model that captures interactions between household heads: development and empirical analysis. *Transportation Research Part B: Methodological*, 36(10):875–896, 2002.

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

Jennifer Tennant, Albert Metz, and Richard Cantor. Moody's financial metrics: Key ratios by rating and industry for global non-financial corporations. Technical report, Moody's Investors Service, December 2007.

Cristiano Varin. On composite marginal likelihoods. *AStA Advances in Statistical Analysis*, 92(1): 1–28, 2008.

Cristiano Varin, Nancy Reid, and David Firth. An overview of composite likelihood methods. *Statistica Sinica*, 21(1):5–42, 2011.

A Simulation study with five outcomes

A simulation study with $J = 5$ outcome variables is performed. The sets of threshold and coefficient parameters are extended for two additional outcomes. For outcome four and five we choose the thresholds $\boldsymbol{\theta}_4 = (-2, -1, 0, 1, 1.5)^\top$ and $\boldsymbol{\theta}_5 = -1.5, -1, -0.5, 0, 0.5, 1, 1.5)^\top$. The following vectors of coefficients are added: $\boldsymbol{\beta}_j = (1.2, -0.2, -1)^\top$, for $j = 4, 5$. We simulate $S = 1000$ data sets with $n = 6000$ subjects. Each subject i has five outcome variables ($J = 5$) yielding in total 30000 observations in the outcome variables. We allow for 6 different sectors with each $n_s = 1000$ subjects and following correlation, matrices:

$$\boldsymbol{R}_1 = \begin{pmatrix} 1.0 & 0.8 & 0.7 & 0.9 & 0.8 \\ 0.8 & 1.0 & 0.8 & 0.8 & 0.7 \\ 0.7 & 0.8 & 1.0 & 0.7 & 0.8 \\ 0.9 & 0.8 & 0.7 & 1.0 & 0.9 \\ 0.8 & 0.7 & 0.8 & 0.9 & 1.0 \end{pmatrix}, \quad \boldsymbol{R}_2 = \begin{pmatrix} 1.0 & 0.4 & 0.5 & 0.6 & 0.5 \\ 0.4 & 1.0 & 0.3 & 0.5 & 0.7 \\ 0.5 & 0.3 & 1.0 & 0.3 & 0.6 \\ 0.6 & 0.5 & 0.3 & 1.0 & 0.5 \\ 0.5 & 0.7 & 0.6 & 0.5 & 1.0 \end{pmatrix}, \quad \boldsymbol{R}_3 = \begin{pmatrix} 1.0 & 0.1 & 0.2 & 0.3 & 0.2 \\ 0.1 & 1.0 & 0.2 & 0.3 & 0.1 \\ 0.2 & 0.2 & 1.0 & 0.1 & 0.3 \\ 0.3 & 0.3 & 0.1 & 1.0 & 0.2 \\ 0.2 & 0.1 & 0.3 & 0.2 & 1.0 \end{pmatrix},$$

$$\boldsymbol{R}_4 = \begin{pmatrix} 1.0 & 0.9 & 0.9 & 0.9 & 0.9 \\ 0.9 & 1.0 & 0.9 & 0.9 & 0.9 \\ 0.9 & 0.9 & 1.0 & 0.9 & 0.9 \\ 0.9 & 0.9 & 0.9 & 1.0 & 0.9 \\ 0.9 & 0.9 & 0.9 & 0.9 & 1.0 \end{pmatrix}, \quad \boldsymbol{R}_5 = \begin{pmatrix} 1.0 & 0.5 & 0.2 & 0.3 & 0.6 \\ 0.5 & 1.0 & 0.2 & 0.3 & 0.1 \\ 0.2 & 0.2 & 1.0 & 0.8 & 0.3 \\ 0.3 & 0.3 & 0.8 & 1.0 & 0.2 \\ 0.6 & 0.1 & 0.3 & 0.2 & 1.0 \end{pmatrix}, \quad \boldsymbol{R}_6 = \begin{pmatrix} 1.0 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 1.0 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 1.0 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 1.0 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.1 & 1.0 \end{pmatrix}.$$

We randomly remove 5% of the first outcome variable, 20% of the second outcome, 50% of the third outcome, 10% of the fourth outcome and 70% of the fifth outcome variable and repeat the simulation.

The findings are similar to the model with three outcome variables. Unreported results show that threshold parameters, coefficients and large correlation parameters are recovered very well for both models. Again, only the estimates of low and moderate correlation parameters suffer in the presence of a high percentage of missing observations. But overall, the model with five different outcome dimension seems to deliver reliable estimates for all parameters. We can conclude that, aside from increasing computation time, increasing number of dimensions in the outcome variables does not pose a problem.

B Tables

Table B.1: This table displays a comparison of pairwise and tripletwise likelihood estimates from the multivariate ordered **probit** using $S = 1000$ simulated data sets, $n = 1000$ subjects and $J = 3$ outcomes.

Parameters		Pairwise Likelihood					Tripletwise Likelihood					Relative Efficiency	
	True Value	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error	Finite Sample Standard Error	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error	Finite Sample Standard Error	$\frac{\text{ASE}_{\text{full}}}{\text{ASE}_{\text{NA}}}$	$\frac{\text{FSSE}_{\text{full}}}{\text{FSSE}_{\text{NA}}}$
$\theta_{1,1}$	−1.0000	−1.0028	0.0028	0.28%	0.0580	0.0577	−1.0025	0.0025	0.25%	0.0585	0.0578	0.9911	0.9995
$\theta_{1,2}$	0.0000	0.0004	−0.0004	–	0.0501	0.0493	0.0003	−0.0003	–	0.0486	0.0492	1.0315	1.0031
$\theta_{1,3}$	1.0000	1.0031	−0.0031	0.31%	0.0580	0.0573	1.0029	−0.0029	0.29%	0.0601	0.0572	0.9652	1.0032
$\theta_{2,1}$	−2.0000	−2.0111	0.0111	0.55%	0.0814	0.0825	−2.0102	0.0102	0.51%	0.0833	0.0817	0.9778	1.0102
$\theta_{2,2}$	0.0000	0.0004	−0.0004	–	0.0496	0.0501	0.0003	−0.0003	–	0.0485	0.0481	1.0223	1.0411
$\theta_{2,3}$	2.0000	2.0115	−0.0115	0.58%	0.0813	0.0803	2.0112	−0.0112	0.56%	0.0840	0.0792	0.9680	1.0141
$\theta_{3,1}$	−1.5000	−1.5060	0.0060	0.40%	0.0658	0.0654	−1.5056	0.0056	0.37%	0.0666	0.0652	0.9882	1.0030
$\theta_{3,2}$	−0.5000	−0.5034	0.0034	0.69%	0.0514	0.0519	−0.5032	0.0032	0.65%	0.0504	0.0515	1.0199	1.0074
$\theta_{3,3}$	0.0000	−0.0004	0.0004	–	0.0496	0.0502	−0.0005	0.0005	–	0.0493	0.0495	1.0055	1.0132
$\theta_{3,4}$	0.5000	0.5010	−0.0010	0.20%	0.0514	0.0517	0.5007	−0.0007	0.14%	0.0517	0.0513	0.9947	1.0078
$\theta_{3,5}$	1.5000	1.5084	−0.0084	0.56%	0.0659	0.0657	1.5081	−0.0081	0.54%	0.0683	0.0659	0.9640	0.9976
$\beta_{1,1}$	1.2000	1.2094	−0.0094	0.78%	0.0531	0.0533	1.2091	−0.0091	0.76%	0.0550	0.0529	0.9653	1.0071
$\beta_{1,2}$	−0.2000	−0.1995	−0.0005	0.23%	0.0389	0.0389	−0.1995	−0.0005	0.25%	0.0411	0.0390	0.9473	0.9984
$\beta_{1,3}$	−1.0000	−1.0040	0.0040	0.40%	0.0490	0.0510	−1.0039	0.0039	0.39%	0.0496	0.0509	0.9881	1.0022
$\beta_{2,1}$	1.2000	1.2111	−0.0111	0.93%	0.0531	0.0524	1.2106	−0.0106	0.89%	0.0550	0.0519	0.9653	1.0095

Continued on next page

Table B.1: (continued)

Parameters		Pairwise Likelihood					Tripletwise Likelihood					Relative Efficiency	
	True Value	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error (Mean)	Finite Sample Standard Error	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error (Median)	Finite Sample Standard Error	$\frac{\text{ASE}_{\text{full}}}{\text{ASE}_{\text{NA}}}$	$\frac{\text{FSSE}_{\text{full}}}{\text{FSSE}_{\text{NA}}}$
$\beta_{2,2}$	−0.2000	−0.2004	0.0004	0.19%	0.0387	0.0381	−0.2002	0.0002	0.09%	0.0408	0.0379	0.9485	1.0063
$\beta_{2,3}$	−1.0000	−1.0044	0.0044	0.44%	0.0489	0.0493	−1.0041	0.0041	0.41%	0.0495	0.0488	0.9881	1.0103
$\beta_{3,1}$	1.2000	1.2098	−0.0098	0.81%	0.0488	0.0480	1.2094	−0.0094	0.78%	0.0509	0.0477	0.9604	1.0066
$\beta_{3,2}$	−0.2000	−0.2006	0.0006	0.30%	0.0364	0.0357	−0.2006	0.0006	0.30%	0.0384	0.0357	0.9483	1.0014
$\beta_{3,3}$	−1.0000	−1.0031	0.0031	0.31%	0.0452	0.0456	−1.0028	0.0028	0.28%	0.0459	0.0457	0.9843	0.9993
ρ_{12}	0.8000	0.8015	−0.0015	0.19%	0.0222	0.0226	0.8017	−0.0017	0.22%	0.0228	0.0219	0.9746	1.0356
ρ_{13}	0.7000	0.6996	0.0004	0.05%	0.0236	0.0236	0.7000	0.0000	0.00%	0.0246	0.0236	0.9580	1.0017
ρ_{23}	0.9000	0.9010	−0.0010	0.11%	0.0129	0.0130	0.9013	−0.0013	0.14%	0.0133	0.0127	0.9669	1.0236

Table B.2: This table displays a comparison of pairwise and tripletwise likelihood estimates from the multivariate ordered **logit** using $S = 1000$ simulated data sets, $n = 1000$ subjects and $J = 3$ outcomes.

Parameters		Pairwise Likelihood					Tripletwise Likelihood					Relative Efficiency	
	True Value	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error	Finite Sample Standard Error	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error	Finite Sample Standard Error	$\frac{\text{ASE}_{\text{full}}}{\text{ASE}_{\text{NA}}}$	$\frac{\text{FSSE}_{\text{full}}}{\text{FSSE}_{\text{NA}}}$
$\theta_{1,1}$	-1.0000	-1.0082	0.0082	0.82%	0.0809	0.0792	-1.0080	0.0080	0.80%	0.0823	0.0793	0.9821	0.9994
$\theta_{1,2}$	0.0000	0.0005	-0.0005	–	0.0734	0.0723	0.0006	-0.0006	–	0.0742	0.0721	0.9897	1.0019
$\theta_{1,3}$	1.0000	1.0096	-0.0096	0.96%	0.0809	0.0786	1.0096	-0.0096	0.96%	0.0864	0.0786	0.9367	0.9997
$\theta_{2,1}$	-2.0000	-2.0160	0.0160	0.80%	0.1002	0.0949	-2.0158	0.0158	0.79%	0.1027	0.0946	0.9764	1.0030
$\theta_{2,2}$	0.0000	-0.0020	0.0020	–	0.0731	0.0717	-0.0016	0.0016	–	0.0739	0.0704	0.9884	1.0188
$\theta_{2,3}$	2.0000	2.0130	-0.0130	0.65%	0.1001	0.0996	2.0129	-0.0129	0.64%	0.1072	0.0986	0.9338	1.0101
$\theta_{3,1}$	-1.5000	-1.5129	0.0129	0.86%	0.0880	0.0821	-1.5122	0.0122	0.81%	0.0912	0.0823	0.9646	0.9982
$\theta_{3,2}$	-0.5000	-0.5082	0.0082	1.64%	0.0743	0.0739	-0.5077	0.0077	1.54%	0.0748	0.0734	0.9936	1.0057
$\theta_{3,3}$	0.0000	-0.0024	0.0024	–	0.0725	0.0713	-0.0020	0.0020	–	0.0737	0.0708	0.9829	1.0081
$\theta_{3,4}$	0.5000	0.5012	-0.0012	0.23%	0.0743	0.0711	0.5014	-0.0014	0.28%	0.0771	0.0706	0.9635	1.0064
$\theta_{3,5}$	1.5000	1.5095	-0.0095	0.63%	0.0881	0.0843	1.5096	-0.0096	0.64%	0.0954	0.0840	0.9232	1.0030
$\beta_{1,1}$	1.2000	1.2098	-0.0098	0.82%	0.0760	0.0735	1.2094	-0.0094	0.78%	0.0812	0.0732	0.9360	1.0038
$\beta_{1,2}$	-0.2000	-0.2043	0.0043	2.15%	0.0630	0.0622	-0.2042	0.0042	2.08%	0.0746	0.0623	0.8445	0.9989
$\beta_{1,3}$	-1.0000	-1.0106	0.0106	1.06%	0.0723	0.0730	-1.0106	0.0106	1.06%	0.0741	0.0728	0.9752	1.0027
$\beta_{2,1}$	1.2000	1.2074	-0.0074	0.62%	0.0729	0.0705	1.2075	-0.0075	0.62%	0.0787	0.0700	0.9262	1.0067
$\beta_{2,2}$	-0.2000	-0.2025	0.0025	1.25%	0.0613	0.0619	-0.2026	0.0026	1.28%	0.0742	0.0621	0.8263	0.9980

Continued on next page

Table B.2: (continued)

Parameters		Pairwise Likelihood					Tripletwise Likelihood					Relative Efficiency	
	True Value	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error (Mean)	Finite Sample Standard Error	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error (Median)	Finite Sample Standard Error	$\frac{\text{ASE}_{\text{full}}}{\text{ASE}_{\text{NA}}}$	$\frac{\text{FSSE}_{\text{full}}}{\text{FSSE}_{\text{NA}}}$
$\beta_{2,3}$	−1.0000	−1.0080	0.0080	0.80%	0.0697	0.0677	−1.0083	0.0083	0.83%	0.0714	0.0673	0.9762	1.0061
$\beta_{3,1}$	1.2000	1.2096	−0.0096	0.80%	0.0720	0.0723	1.2095	−0.0095	0.79%	0.0774	0.0720	0.9294	1.0037
$\beta_{3,2}$	−0.2000	−0.2040	0.0040	2.00%	0.0602	0.0613	−0.2039	0.0039	1.94%	0.0730	0.0612	0.8247	1.0010
$\beta_{3,3}$	−1.0000	−1.0113	0.0113	1.13%	0.0685	0.0684	−1.0114	0.0114	1.14%	0.0703	0.0681	0.9739	1.0048
ρ_{12}	0.8000	0.7997	0.0003	0.04%	0.0190	0.0189	0.7998	0.0002	0.02%	0.0203	0.0190	0.9376	0.9967
ρ_{13}	0.7000	0.6988	0.0012	0.17%	0.0240	0.0242	0.6989	0.0011	0.16%	0.0256	0.0240	0.9382	1.0097
ρ_{23}	0.9000	0.9003	−0.0003	0.03%	0.0107	0.0105	0.9004	−0.0004	0.04%	0.0116	0.0105	0.9200	1.0037

Table B.3: This table displays a comparison of the full observations model and the missing observations model for pairwise likelihood threshold parameter estimates as well as coefficient estimates from the multivariate ordered response model with **probit** link using the $S = 1000$ simulated data sets, $n_s = 1000$ subjects for each sector and $J = 3$ outcome dimensions.

Parameters		Full Observations Model					Missing Observations Model					Relative Efficiency	
	True Value	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error	Finite Sample Standard Error	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error	Finite Sample Standard Error	$\frac{\text{ASE}_{\text{full}}}{\text{ASE}_{\text{NA}}}$	$\frac{\text{FSSE}_{\text{full}}}{\text{FSSE}_{\text{NA}}}$
$\theta_{1,1}$	−1.0000	−1.0020	0.0020	0.20%	0.0227	0.0287	−1.0022	0.0022	0.22%	0.0250	0.0308	0.9076	0.9331
$\theta_{1,2}$	0.0000	0.0001	−0.0001	–	0.0194	0.0224	0.0005	−0.0005	–	0.0215	0.0239	0.9019	0.9364
$\theta_{1,3}$	1.0000	1.0015	−0.0015	0.15%	0.0226	0.0281	1.0018	−0.0018	0.18%	0.0249	0.0294	0.9076	0.9544
$\theta_{2,1}$	−2.0000	−2.0041	0.0041	0.21%	0.0323	0.0494	−2.0050	0.0050	0.25%	0.0383	0.0542	0.8425	0.9110
$\theta_{2,2}$	0.0000	−0.0005	0.0005	–	0.0193	0.0255	−0.0000	0.0000	–	0.0228	0.0283	0.8438	0.9011
$\theta_{2,3}$	2.0000	2.0040	−0.0040	0.20%	0.0322	0.0486	2.0044	−0.0044	0.22%	0.0382	0.0524	0.8429	0.9267
$\theta_{3,1}$	−1.5000	−1.5013	0.0013	0.09%	0.0256	0.0366	−1.5007	0.0007	0.05%	0.0365	0.0455	0.7012	0.8044
$\theta_{3,2}$	−0.5000	−0.5003	0.0003	0.07%	0.0200	0.0253	−0.4999	−0.0001	0.03%	0.0282	0.0311	0.7075	0.8138
$\theta_{3,3}$	0.0000	0.0001	−0.0001	–	0.0191	0.0226	0.0002	−0.0002	–	0.0269	0.0293	0.7108	0.7699
$\theta_{3,4}$	0.5000	0.4995	0.0005	0.10%	0.0200	0.0249	0.5000	0.0000	0.00%	0.0282	0.0316	0.7079	0.7869
$\theta_{3,5}$	1.5000	1.5039	−0.0039	0.26%	0.0256	0.0370	1.5042	−0.0042	0.28%	0.0366	0.0462	0.7006	0.8014
$\beta_{1,1}$	1.2000	1.2033	−0.0033	0.28%	0.0204	0.0265	1.2039	−0.0039	0.32%	0.0227	0.0277	0.9027	0.9569
$\beta_{1,2}$	−0.2000	−0.2004	0.0004	0.22%	0.0149	0.0156	−0.2007	0.0007	0.33%	0.0165	0.0169	0.8987	0.9260
$\beta_{1,3}$	−1.0000	−1.0019	0.0019	0.19%	0.0188	0.0244	−1.0027	0.0027	0.27%	0.0209	0.0260	0.9025	0.9383
$\beta_{2,1}$	1.2000	1.2036	−0.0036	0.30%	0.0206	0.0286	1.2040	−0.0040	0.34%	0.0244	0.0313	0.8437	0.9143

Continued on next page

Table B.3: (continued)

Parameters		Full Observations Model					Missing Observations Model					Relative Efficiency	
	True Value	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error (Mean)	Finite Sample Standard Error	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error (Median)	Finite Sample Standard Error	$\frac{\text{ASE}_{\text{full}}}{\text{ASE}_{\text{NA}}}$	$\frac{\text{FSSE}_{\text{full}}}{\text{FSSE}_{\text{NA}}}$
$\beta_{2,2}$	-0.2000	-0.2002	0.0002	0.12%	0.0148	0.0174	-0.2001	0.0001	0.03%	0.0175	0.0204	0.8455	0.8555
$\beta_{2,3}$	-1.0000	-1.0014	0.0014	0.14%	0.0189	0.0267	-1.0012	0.0012	0.12%	0.0224	0.0298	0.8447	0.8979
$\beta_{3,1}$	1.2000	1.2027	-0.0027	0.22%	0.0187	0.0265	1.2027	-0.0027	0.23%	0.0262	0.0323	0.7123	0.8199
$\beta_{3,2}$	-0.2000	-0.1999	-0.0001	0.04%	0.0138	0.0143	-0.1993	-0.0007	0.36%	0.0191	0.0193	0.7224	0.7409
$\beta_{3,3}$	-1.0000	-1.0011	0.0011	0.11%	0.0173	0.0233	-1.0017	0.0017	0.17%	0.0242	0.0289	0.7145	0.8061
ρ^1_{12}	0.8000	0.8006	-0.0006	0.08%	0.0211	0.0218	0.8007	-0.0007	0.08%	0.0240	0.0244	0.8800	0.8927
ρ^1_{13}	0.7000	0.6970	0.0030	0.43%	0.0216	0.0221	0.6990	0.0010	0.14%	0.0303	0.0312	0.7147	0.7071
ρ^1_{23}	0.9000	0.8996	0.0004	0.04%	0.0122	0.0124	0.9008	-0.0008	0.09%	0.0181	0.0182	0.6720	0.6830
ρ^2_{12}	0.5000	0.4988	0.0012	0.25%	0.0355	0.0359	0.4995	0.0005	0.11%	0.0411	0.0412	0.8633	0.8723
ρ^2_{13}	0.3000	0.2974	0.0026	0.87%	0.0381	0.0411	0.2971	0.0029	0.98%	0.0555	0.0578	0.6872	0.7114
ρ^2_{23}	0.4000	0.3984	0.0016	0.39%	0.0360	0.0364	0.3976	0.0024	0.59%	0.0587	0.0607	0.6138	0.6005
ρ^3_{12}	0.2000	0.1979	0.0021	1.04%	0.0437	0.0455	0.1991	0.0009	0.43%	0.0503	0.0526	0.8686	0.8645
ρ^3_{13}	0.3000	0.2981	0.0019	0.64%	0.0382	0.0406	0.2959	0.0041	1.35%	0.0544	0.0568	0.7017	0.7147
ρ^3_{23}	0.1000	0.0981	0.0019	1.90%	0.0419	0.0415	0.0972	0.0028	2.83%	0.0644	0.0660	0.6505	0.6293
ρ^4_{12}	0.9000	0.9015	-0.0015	0.16%	0.0151	0.0160	0.9017	-0.0017	0.19%	0.0172	0.0187	0.8772	0.8556
ρ^4_{13}	0.9000	0.9002	-0.0002	0.02%	0.0097	0.0100	0.9003	-0.0003	0.03%	0.0139	0.0147	0.6942	0.6785
ρ^4_{23}	0.9000	0.9003	-0.0003	0.03%	0.0121	0.0124	0.9008	-0.0008	0.09%	0.0190	0.0190	0.6375	0.6549
ρ^5_{12}	0.8000	0.7999	0.0001	0.01%	0.0212	0.0218	0.8003	-0.0003	0.04%	0.0241	0.0248	0.8784	0.8795

Continued on next page

Table B.3: (continued)

Parameters		Full Observations Model					Missing Observations Model					Relative Efficiency	
	True Value	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error (Mean)	Finite Sample Standard Error	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error (Median)	Finite Sample Standard Error	$\frac{\text{ASE}_{\text{full}}}{\text{ASE}_{\text{NA}}}$	$\frac{\text{FSSE}_{\text{full}}}{\text{FSSE}_{\text{NA}}}$
ρ_{13}^{5}	0.3000	0.2976	0.0024	0.81%	0.0381	0.0411	0.2970	0.0030	0.99%	0.0570	0.0579	0.6688	0.7104
ρ_{23}^{5}	0.6000	0.6003	−0.0003	0.06%	0.0284	0.0288	0.6004	−0.0004	0.07%	0.0452	0.0433	0.6296	0.6640
ρ_{12}^{6}	0.1000	0.1011	−0.0011	1.11%	0.0448	0.0472	0.1008	−0.0008	0.77%	0.0509	0.0524	0.8807	0.9008
ρ_{13}^{6}	0.1000	0.1012	−0.0012	1.23%	0.0416	0.0409	0.1023	−0.0023	2.26%	0.0604	0.0599	0.6894	0.6835
ρ_{23}^{6}	0.1000	0.1005	−0.0005	0.45%	0.0417	0.0426	0.0986	0.0014	1.41%	0.0671	0.0681	0.6223	0.6255

Table B.4: This table displays a comparison of the full observations model and the missing observations model for pairwise likelihood threshold parameter estimates as well as coefficient estimates from the multivariate ordered response model with **logit** link using the $S = 1000$ simulated data sets, $n_s = 1000$ subjects for each sector and $J = 3$ outcome dimensions.

Parameters		Full Observations Model					Missing Observations Model					Relative Efficiency	
	True Value	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error (Mean)	Finite Sample Standard Error	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error (Median)	Finite Sample Standard Error	$\frac{\text{ASE}_{\text{full}}}{\text{ASE}_{\text{NA}}}$	$\frac{\text{FSSE}_{\text{full}}}{\text{FSSE}_{\text{NA}}}$
$\theta_{1,1}$	−1.0000	−1.0034	0.0034	0.34%	0.0309	0.0308	−1.0035	0.0035	0.35%	0.0343	0.0333	0.9024	0.9266
$\theta_{1,2}$	0.0000	−0.0006	0.0006	–	0.0280	0.0276	−0.0006	0.0006	–	0.0312	0.0308	0.8994	0.8960
$\theta_{1,3}$	1.0000	1.0032	−0.0032	0.32%	0.0309	0.0293	1.0033	−0.0033	0.33%	0.0343	0.0323	0.9020	0.9068
$\theta_{2,1}$	−2.0000	−2.0020	0.0020	0.10%	0.0385	0.0385	−2.0030	0.0030	0.15%	0.0457	0.0458	0.8435	0.8417
$\theta_{2,2}$	0.0000	−0.0001	0.0001	–	0.0280	0.0271	−0.0006	0.0006	–	0.0331	0.0323	0.8442	0.8370
$\theta_{2,3}$	2.0000	2.0023	−0.0023	0.11%	0.0386	0.0374	2.0035	−0.0035	0.17%	0.0457	0.0428	0.8434	0.8750
$\theta_{3,1}$	−1.5000	−1.5063	0.0063	0.42%	0.0337	0.0335	−1.5066	0.0066	0.44%	0.0475	0.0478	0.7097	0.7014
$\theta_{3,2}$	−0.5000	−0.5025	0.0025	0.50%	0.0284	0.0280	−0.5037	0.0037	0.73%	0.0398	0.0398	0.7143	0.7038
$\theta_{3,3}$	0.0000	−0.0005	0.0005	–	0.0277	0.0263	−0.0010	0.0010	–	0.0387	0.0379	0.7152	0.6951
$\theta_{3,4}$	0.5000	0.5024	−0.0024	0.48%	0.0284	0.0279	0.5016	−0.0016	0.33%	0.0398	0.0389	0.7142	0.7174
$\theta_{3,5}$	1.5000	1.5043	−0.0043	0.29%	0.0337	0.0338	1.5031	−0.0031	0.21%	0.0475	0.0471	0.7097	0.7169
$\beta_{1,1}$	1.2000	1.2048	−0.0048	0.40%	0.0287	0.0283	1.2050	−0.0050	0.42%	0.0320	0.0321	0.8958	0.8791
$\beta_{1,2}$	−0.2000	−0.2009	0.0009	0.47%	0.0236	0.0232	−0.2004	0.0004	0.21%	0.0265	0.0263	0.8932	0.8815
$\beta_{1,3}$	−1.0000	−1.0049	0.0049	0.49%	0.0272	0.0271	−1.0050	0.0050	0.50%	0.0304	0.0294	0.8949	0.9210
$\beta_{2,1}$	1.2000	1.2018	−0.0018	0.15%	0.0273	0.0270	1.2025	−0.0025	0.21%	0.0323	0.0320	0.8446	0.8452

Continued on next page

Table B.4: (continued)

Parameters		Full Observations Model					Missing Observations Model					Relative Efficiency	
	True Value	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error	Finite Sample Standard Error	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error	Finite Sample Standard Error	$\frac{\text{ASE}_{\text{full}}}{\text{ASE}_{\text{NA}}}$	$\frac{\text{FSSE}_{\text{full}}}{\text{FSSE}_{\text{NA}}}$
$\beta_{2,2}$	−0.2000	−0.2006	0.0006	0.32%	0.0228	0.0229	−0.2003	0.0003	0.14%	0.0270	0.0270	0.8448	0.8469
$\beta_{2,3}$	−1.0000	−1.0017	0.0017	0.17%	0.0260	0.0257	−1.0019	0.0019	0.19%	0.0308	0.0293	0.8443	0.8779
$\beta_{3,1}$	1.2000	1.2048	−0.0048	0.40%	0.0269	0.0271	1.2059	−0.0059	0.49%	0.0371	0.0376	0.7239	0.7194
$\beta_{3,2}$	−0.2000	−0.2008	0.0008	0.40%	0.0223	0.0215	−0.2015	0.0015	0.73%	0.0306	0.0311	0.7298	0.6924
$\beta_{3,3}$	−1.0000	−1.0043	0.0043	0.43%	0.0255	0.0258	−1.0053	0.0053	0.53%	0.0353	0.0355	0.7239	0.7256
ρ_{12}^{1}	0.8000	0.8003	−0.0003	0.04%	0.0174	0.0173	0.8008	−0.0008	0.10%	0.0198	0.0200	0.8807	0.8612
ρ_{13}^{1}	0.7000	0.7002	−0.0002	0.03%	0.0218	0.0221	0.6995	0.0005	0.07%	0.0307	0.0321	0.7105	0.6894
ρ_{23}^{1}	0.9000	0.9002	−0.0002	0.02%	0.0097	0.0097	0.8995	0.0005	0.06%	0.0147	0.0149	0.6632	0.6501
ρ_{12}^{2}	0.5000	0.5006	−0.0006	0.11%	0.0326	0.0317	0.4998	0.0002	0.04%	0.0378	0.0372	0.8610	0.8538
ρ_{13}^{2}	0.3000	0.3005	−0.0005	0.15%	0.0386	0.0367	0.3013	−0.0013	0.44%	0.0563	0.0545	0.6853	0.6739
ρ_{23}^{2}	0.4000	0.3979	0.0021	0.53%	0.0344	0.0334	0.3957	0.0043	1.07%	0.0563	0.0562	0.6111	0.5952
ρ_{12}^{3}	0.2000	0.1987	0.0013	0.65%	0.0412	0.0412	0.1999	0.0001	0.04%	0.0474	0.0475	0.8694	0.8667
ρ_{13}^{3}	0.3000	0.2990	0.0010	0.33%	0.0386	0.0379	0.3004	−0.0004	0.13%	0.0547	0.0534	0.7050	0.7101
ρ_{23}^{3}	0.1000	0.0980	0.0020	1.96%	0.0405	0.0406	0.0963	0.0037	3.74%	0.0625	0.0633	0.6482	0.6412
ρ_{12}^{4}	0.9000	0.9003	−0.0003	0.04%	0.0109	0.0110	0.9007	−0.0007	0.08%	0.0125	0.0128	0.8679	0.8608
ρ_{13}^{4}	0.9000	0.9005	−0.0005	0.06%	0.0089	0.0085	0.9008	−0.0008	0.08%	0.0129	0.0127	0.6933	0.6707
ρ_{23}^{4}	0.9000	0.9002	−0.0002	0.02%	0.0097	0.0100	0.9011	−0.0011	0.12%	0.0152	0.0156	0.6376	0.6397
ρ_{12}^{5}	0.8000	0.7992	0.0008	0.10%	0.0175	0.0176	0.7997	0.0003	0.04%	0.0199	0.0204	0.8767	0.8620

Continued on next page

Table B.4: (continued)

Parameters		Full Observations Model					Missing Observations Model					Relative Efficiency	
	True Value	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error	Finite Sample Standard Error	Mean Estimate	Bias	Absolute Perc. Bias	Asympt. Standard Error	Finite Sample Standard Error	$\frac{\mathrm{ASE}_{\mathrm{full}}}{\mathrm{ASE}_{\mathrm{NA}}}$	$\frac{\mathrm{FSSE}_{\mathrm{full}}}{\mathrm{FSSE}_{\mathrm{NA}}}$
ρ_{13}^{5}	0.3000	0.3009	−0.0009	0.30%	0.0385	0.0383	0.3012	−0.0012	0.41%	0.0576	0.0573	0.6688	0.6675
ρ_{23}^{5}	0.6000	0.6011	−0.0011	0.18%	0.0264	0.0266	0.6017	−0.0017	0.28%	0.0420	0.0417	0.6295	0.6382
ρ_{12}^{6}	0.1000	0.1003	−0.0003	0.28%	0.0425	0.0424	0.1017	−0.0017	1.65%	0.0482	0.0485	0.8812	0.8745
ρ_{13}^{6}	0.1000	0.1025	−0.0025	2.49%	0.0421	0.0403	0.1064	−0.0064	6.35%	0.0610	0.0621	0.6907	0.6491
ρ_{23}^{6}	0.1000	0.1006	−0.0006	0.59%	0.0405	0.0402	0.1020	−0.0020	2.03%	0.0649	0.0662	0.6242	0.6070

Table B.5: This table displays the threshold parameter estimates from the multivariate ordered **probit** model using the multiple corporate credit ratings data set.

Probit link							
	S&P		Fitch			Moody's	
Thresholds	Est.	SE	Est.	SE	Thresholds	Est.	SE
					Ca|Caa	−4.1522	0.0736
CCC/C|B	−3.5216	0.0515	−3.1305	0.1003	Caa|B	−2.6254	0.0474
B|BB	−1.4999	0.0371	−1.6147	0.0689	B|Ba	−1.0285	0.0390
BB|BBB	−0.3165	0.0359	−0.5537	0.0629	Ba|Baa	−0.2359	0.0386
BBB|A	1.0612	0.0377	0.8229	0.0632	Baa|A	1.1502	0.0406
A|AA	2.4967	0.0455	2.3785	0.0723	A|Aa	2.6008	0.0508
AA|AAA	3.5155	0.0598	3.6558	0.1045	Aa|Aaa	3.6487	0.0667

Table B.6: This table displays the regression coefficients from the multivariate ordered **probit** model using the multiple corporate credit ratings data set.

	S&P		Moody's		Fitch	
Covariate	Est.	SE	Est.	SE	Est.	SE
operating CF cov.	0.0551	0.0123	0.0695	0.0134	0.0591	0.0181
cash/assets	−0.0598	0.0099	−0.0488	0.0107	−0.0778	0.0139
tangibility	0.1337	0.0120	0.1504	0.0132	0.1154	0.0172
debt/assets	−0.3654	0.0131	−0.3390	0.0144	−0.4242	0.0230
ST debt/debt	0.1010	0.0108	0.1178	0.0121	0.1208	0.0163
ret.earnings/assets	0.4105	0.0124	0.3937	0.0141	0.3587	0.0175
return on capital	0.1795	0.0132	0.1911	0.0141	0.1893	0.0181
EBIT margin	0.1193	0.0120	0.1064	0.0121	0.1235	0.0154
R&D/assets	0.1583	0.0098	0.1453	0.0107	0.1642	0.0137
capex/assets	−0.0518	0.0119	−0.0698	0.0128	−0.0194	0.0197
RSIZE	0.5018	0.0123	0.5446	0.0138	0.4276	0.0169
BETA	−0.1139	0.0101	−0.1035	0.0108	−0.1252	0.0151
SIGMA	−0.3514	0.0158	−0.3184	0.0170	−0.3031	0.0289
MB	−0.1230	0.0111	−0.1084	0.0129	−0.0619	0.0167