Journal of Empirical Finance 69 (2022) 224–240

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

A corporate credit rating model with autoregressive errors

Rainer Hirk [a,*,1], Laura Vana [b,2], Kurt Hornik [a,3]

[a] Vienna University of Economics and Business, Institute for Statistics and Mathematics, Welthandelsplatz 1, 1020 Vienna, Austria
[b] Vienna University of Technology, Institute of Statistics and Mathematical Methods in Economics, Austria

ARTICLE INFO

Keywords:
Composite likelihood
Corporate credit ratings
Longitudinal ordinal regression model
Ordinal LASSO
Predictive performance

ABSTRACT

In this paper we propose a longitudinal credit rating model which accounts for the serial correlation in the ratings. We achieve this by imposing an autoregressive structure of order one on the errors of a multivariate ordinal regression model. The longitudinal structure of the model improves significantly both the goodness-of-fit and predictive performance compared to static models. By modeling the joint distribution of the ratings over time, the framework allows us to obtain predictions conditional on the past rating history of a firm, which clearly outperform the unconditional predictions both in- and out-of-sample. This shows the importance of incorporating past rating information in the prediction. Another upside lies in the framework's ability to deal with missing rating observations. A real data example is provided by using a sample of US publicly traded corporates rated by S&P for the years 1985–2016. The determinants of corporate credit ratings are pre-selected using the ordinal version of the least absolute shrinkage and selection operator (LASSO). Additionally, as a model extension we allow the regression coefficients of the selected variables to vary over time in the longitudinal model. This allows us to gain a better understanding of the drivers and evolution of the rating behavior over the sample period. Finally, based on the longitudinal model with LASSO selected variables, we find evidence that S&P exhibits procyclical aspects in their rating behavior.

1. Introduction

The last decades have witnessed an increased interest from practitioners, researchers and regulators alike in developing tools for appropriately understanding and modeling credit ratings from the large credit rating agencies (CRAs). During the financial crisis of 2007–2009 the CRAs came under special scrutiny after failing to assess credit risk appropriately. They, however, continue to be key players in financial markets, with their ratings being the most common and widely used measure of corporate credit quality (Hilscher and Wilson, 2016), which in turn impacts a firm's cost of funding as well as the firm's growth (Baghai et al., 2014). The modeling of credit ratings is therefore relevant not only for prediction purposes but also to shed light on the behavior and methodology of the CRAs. The literature concerned with the prediction of ratings provided by the big CRAs focuses in general on building econometric models which rely on firm characteristics as control variables and drivers of credit quality. These models quantify empirically the importance of the drivers in explaining (and predicting) the observed credit ratings. The majority of the approaches consider single-period data or assume independent rating observations over time (e.g., Blume et al., 1998; Alp, 2013; Baghai et al., 2014). However, the importance of accounting for the correlation in the credit ratings over time has been documented in the literature (e.g., Altman and Kao, 1992; Hwang, 2013; Dimitrakopoulos and Kolossiatis, 2016; Tuzcuoğlu, 2019). Furthermore, especially at a firm-level,

* Corresponding author.
E-mail addresses: rainer.hirk@wu.ac.at (R. Hirk), laura.vana@wu.ac.at (L. Vana), kurt.hornik@wu.ac.at (K. Hornik).

1 Postdoctoral researcher.
2 Assistant Professor.
3 Professor.

https://doi.org/10.1016/j.jempfin.2022.09.002
Received 27 July 2021; Received in revised form 5 September 2022; Accepted 10 September 2022
Available online 22 September 2022
0927-5398/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

the choice of variables to include in a rating model is of particular importance in the model building phase, but in the absence of a clear theory which can inform researchers on the determinants of credit quality, the choice of the firm characteristics to be included in such econometric models is ambiguous. While a consensus exists in the literature that key financial aspects of a firm such as profitability, leverage, liquidity have a significant impact on the credit risk measures, variable selection in a credit risk context has primarily relied on expert judgment and statistical approaches to variable selection have been rather limited (see e.g., Tian et al., 2015; Sermpinis et al., 2018).

In this paper we propose a modeling framework for S&P corporate ratings which extends the approaches in the literature along two directions. On the one hand, we tackle the lack of consensus on which variables to include in a corporate credit rating model by performing an empirical variable selection exercise using the LASSO (least absolute shrinkage and selection operator of Tibshirani, 1996). The variables are selected by the LASSO procedure to maximize the model's ability to predict out-of-sample (more specifically out-of-time and out-of-firm), rather than in-sample. On the other hand, we account for the dependence in the ratings over time in the modeling framework by leveraging the flexibility of the class of multivariate ordinal regression models (see, e.g., Hirk et al., 2020). We build a longitudinal model which accounts for the persistence in ratings through firm error terms with an autoregressive structure and takes into account the ordinal nature of credit ratings. The use of an autoregressive (i.e., dynamic) specification in the error terms allows us to account for dependence on past ratings and for heterogeneity not captured by the financial statement covariates and also alleviates the omitted variable problem. Examples of typically unobserved factors include management ability, competition, or aspects related to human resources. The ordinal nature of the credit ratings is reflected in the model by assuming that the observed ordinal ratings are coarser versions of underlying latent continuous variables. This implies that the CRAs compute a creditworthiness score on a numeric scale by employing a statistical model and then obtain the final published credit ratings by slotting the underlying score into classes using appropriate thresholds.

The proposed framework has several features which are beneficial for the use-case of rating prediction: (i) Given that the panel of corporate credit ratings is unbalanced, with ratings being (temporarily) withdrawn before the end of the sample period, having a framework which can deal with missing rating observations is desirable. We model the repeated S&P ratings over time as responses following a multivariate distribution. This allows us to easily deal with missing observations and also to generate prediction and evaluation measures which rely on the multivariate structure. (ii) More specifically, we propose to evaluate the model based on its ability to predict the ratings at each time point, conditional on the observed past rating history. These conditional predictions are straightforward to compute from the multivariate distribution. By not including the past rating history in the model directly as covariates, but rather as responses, we also alleviate the issue of how to impute missing ratings on the covariate side. (iii) Another novel feature in the model is the use of the multivariate logistic error distribution (O'Brien and Dunson, 2004) which has the advantage that the regression coefficients can be interpreted as log odds ratios and also shows a better goodness-of-fit and predictive performance when compared to the more commonly used multivariate normal distribution. One should note that this distribution has heavier tails than the multivariate normal distribution.

We employ the proposed framework on a sample of publicly rated US companies which get at least one rating from S&P over the period 1985–2016. Firstly, by using the ordinal version of the LASSO, we provide guidance on the importance of an extensive collection of 78 market and financial ratios and pre-select 13 variables for inclusion in the model. The ratios selected by the LASSO are to a large extent in line with existing literature and measure different aspects of firm performance such as size (2 ratios), profitability (2 ratios), liquidity (1 ratio), leverage and capital structure (5 ratios), market-based systematic and idiosyncratic risk (2 ratios) and whether the company is a dividend payer. While the majority of these ratios are the same as the ones employed by Alp (2013), we observe that ratios related to interest coverage, tangibility and cash-flow are not selected by the empirical variable selection exercise. Also, our model includes more ratios related to leverage and capital structure than alternative models. This hints towards the fact that measuring different aspects of leverage is important for the modeling context. Secondly, by employing the dynamic ordinal regression model, we achieve an improvement in both the goodness-of-fit and predictive performance measures which include past rating information when compared to the model's static counterpart. When comparing the model containing the variables selected by the LASSO with models containing the ratios in Alp (2013) and Baghai et al. (2014) we see a superior performance of the former. However, it is to be noted that the improvement of the LASSO exercise over the model with Alp (2013) ratios is rather small. Therefore, we conclude that the relative importance of adding the longitudinal structure in the model specification is larger than the improvement gained by the variable selection exercise over the variables employed in Alp (2013). We also illustrate how the framework can be used to estimate a time-varying coefficients model which can provide insights into the evolution of the rating behavior over the sample period. For some variables, such as retained earnings to assets, we observe a decrease in the relevance over the sample period in explaining the S&P ratings, while ratios such as convertible debt to assets have increasing explanatory power after the financial crisis of 2007–2009. Finally, based on the longitudinal model with LASSO selected variables, we find evidence that S&P exhibits procyclical aspects in their rating behavior.

Historically, various modeling approaches have been employed for building predictive models of credit ratings including regression analysis (where the ordinal credit ratings are transformed to a numeric scale, e.g. Horrigan, 1966; Pogue and Soldofsky, 1969; West, 1970), discriminant analysis (Pinches and Mingo, 1973; Altman and Katz, 1976) and ordinal regression models with probit and logit links (Kaplan and Urwitz, 1979; Ederington, 1985; Gentry et al., 1988; Blume et al., 1998; Poon et al., 1999; Alp, 2013). More complex approaches have also been employed in e.g., Hwang et al. (2010), who replace the linear regression function in the usual ordinal probit regression model with a more flexible semi-parametric function. These approaches either consider single-period data or they ignore the longitudinal structure of multi-period data by building static models which assume independent observations over time. The use of panel data models for credit rating prediction is rather limited in the literature, likely due to the computational complexity introduced by the modeling of the multivariate structure of the ratings and the scarcity of software tools to estimate such models. In order to account for the history of past ratings in a static panel data model, some authors include previous ratings as covariates in the regression function (e.g., Sermpinis et al., 2018). As previously mentioned, this approach has limitations in the case of missing ratings as it is not straightforward how to handle missing values in the covariates. Sermpinis et al. (2018) also select relevant variables in modeling market implied ratings using the LASSO and show improved predictive power, even though their proposed models still contain a larger number of variables than standard models in the literature (e.g., Alp, 2013; Baghai et al., 2014). Several papers account for the fact that the company ratings observed repeatedly over time are correlated (the so-called dynamic setting) and show improved predictive performance. A random effects model is proposed by Afonso et al. (2009), who include firm random effects to account for the fact that sovereign ratings are likely to be correlated over time. Hwang (2013) propose a similar model using generalized estimating equations. Both approaches assume an exchangeable correlation structure on the time series of ratings which implies a constant dependence regardless of the time lag, an assumption which is rather unrealistic. Dynamic specifications with an autoregressive structure, where the correlation decreases exponentially with the time lag, have been employed in the literature in different contexts. For example, Creal et al. (2014) build a dynamic factor model for modeling Moody's rating transitions together with macroeconomic and loss given default data, but the setting is applicable at a portfolio-level, not firm-level. For sovereign credit ratings, dynamic ordinal models are employed in Dimitrakopoulos and Kolossiatis (2016), Reusens and Croux (2017). For corporate ratings, the framework proposed in this paper is closest to Hirk et al. (2019) and Tuzcuoğlu (2019) in that it employs a multivariate ordinal model, but enhances the modeling framework in several directions. Firstly, none of these two papers addressed the issue of variable selection. Hirk et al. (2019) applied a cross-sectional multivariate ordinal model to investigate the correlation between the ratings assigned by S&P, Moody's and Fitch. This model is static, in that it does not account for the serial correlation of the ratings over time. Through the use of the conditional predictions, this cross-sectional model can be evaluated based on its ability to predict the S&P rating at a point in time, conditional on knowing ratings from the other rating agencies at the same time-point (if such ratings are available). If there is no information on ratings from other CRAs, the cross-sectional model would deliver the unconditional (i.e., marginal) rating predictions. These predictions can then be used to compare the model in Hirk et al. (2019) with the model proposed in this work. Tuzcuoğlu (2019) builds a longitudinal multivariate probit model with random effects and illustrates its use on a data set of corporate bond ratings. In contrast to this author's approach, we propose to use a multivariate logistic distribution for the errors, and propose the use of predictions conditional on the past rating observations for model evaluation.

The remainder of this paper is structured as follows: Section 2 introduces the empirical strategy which relies on the ordinal regression modeling framework. The data together with the variable selection procedure is presented in Section 3. Section 4 presents and discusses the results of the longitudinal model. Section 5 concludes.

2. Methodology

The model class of multivariate ordinal regression models offers a framework where the time dependence of ratings observed repeatedly over time for one company can be modeled by imposing an autoregressive structure of order one ($AR(1)$) on the error terms of the model. Such a longitudinal model has the advantage that the dependence of a rating on the ratings from the previous years can readily be accounted for in the model. This is in contrast to single-year or static models where the ratings of the same company are assumed to be independent.

Let $i = 1, \dots, N$ denote the firm index and $t \in T_i$ the year index of one observation in the sample. For each firm i, T_i is the subset of all available time points in the sample $\{1, \dots, T\}$ for which a rating is observed and $q_i = |T_i|$ is the cardinality of the set T_i. A multivariate ordinal regression model with $AR(1)$ errors is employed to predict the behavior of the ordinal variable Y which can take one of $\{1, \dots, K\}$ ordered categories with a set of p predictors for each point in time. Assuming that the observed ordinal variable Y is a categorized version of an underlying variable $\widetilde{Y}$, which in our application can be interpreted as a latent creditworthiness score, Y and $\widetilde{Y}$ are linked through the parameters θ_r, which serve as "slotting" threshold parameters. A regression model is then assumed on the underlying variable $\widetilde{Y}$:

$$Y_{i,t} = r_{i,t} \iff \theta_{r_{i,t}-1} < \widetilde{Y}_{i,t} \leq \theta_{r_{i,t}}, \qquad \widetilde{Y}_{i,t} = \boldsymbol{x}_{i,t}^\top \boldsymbol{\beta} + \epsilon_{i,t}, \tag{1}$$

where $r_{i,t}$ is the rating observed for firm i at time t, $\boldsymbol{x}_{i,t}$ is a $(p \times 1)$ vector containing the p predictors for firm i in year t, $\boldsymbol{\beta}$ is a $(p \times 1)$ vector of regression coefficients, θ_r are threshold parameters for each ordered class r out of the K classes which satisfy the restriction $-\infty = \theta_0 < \theta_1 < \cdots < \theta_K = \infty$ and $\epsilon_{i,t}$ is an error term. The multivariate structure is obtained by assuming that ϵ_i, the q_i-dimensional vector of error terms corresponding to firm i, follows a multivariate distribution. In this paper we consider the multivariate logistic distribution which is built by using a multivariate t-copula for the dependence structure across time and has logistic margins (O'Brien and Dunson, 2004). For the t-copula we assume mean zero and correlation matrix $\boldsymbol{C}$. The $AR(1)$ structure of the errors is reflected in the entries of $\boldsymbol{C}$: $\mathrm{cor}(\epsilon_{i,k}, \epsilon_{i,l}) = c_{k,l} = \rho^{|k-l|}$ where k and l represent time points. Alternatively, another typical choice for the multivariate distribution function is given by the multivariate normal distribution. One can choose the most appropriate multivariate distribution function by relying on e.g., the Akaike or the Bayesian information criterion.

For each firm i, the log-likelihood of the longitudinal model is given by the log-probability of events $Y_{i,t} = r_{i,t}$ for all $t \in T_i$:

$$\ell^{\mathrm{AR}(1)}(\boldsymbol{\theta}, \boldsymbol{\beta}, \rho) = \sum_{i=1}^{N} \log \mathbb{P}\left(\bigcap_{t \in T_i} Y_{i,t} = r_{i,t} \right). \tag{2}$$

The estimation is performed using composite maximum likelihood methods (see Varin, 2008, for an overview), by replacing the original likelihood, which involves the computation of a q_i-dimensional integral for each i, by the product of lower dimensional probabilities. The pairwise log-likelihood is composed of all bivariate log-probabilities corresponding to pairs of available ratings:

$$p\ell(\boldsymbol{\theta}, \boldsymbol{\beta}, \rho) = \sum_{i=1}^{N} \sum_{\substack{s<t,\\ s,t\in T_i}} \log \mathbb{P}(Y_{i,s} = r_{i,s}, Y_{i,t} = r_{i,t}).$$

As absolute scale and location are not identifiable in an ordinal model we fix the marginal variance of the underlying process $\tilde{Y}$ to one and the intercept in the latent regression model in Eq. (1) to zero. When applying composite likelihood estimation, composite likelihood information criteria (e.g., CLAIC, CLBIC), which adapt the classical measures to account for the misspecified likelihood, can be used for model comparison (Varin and Vidoni, 2005).

One major advantage of the proposed modeling class is that it can easily be applied to settings where missing observations are present, i.e., there exist firms for which $T_i \neq \{1, \ldots, T\}$. Another advantage of employing a multivariate structure on the errors is that one can estimate the probability of observing a certain rating over the next period conditional on the ratings of the past periods. Given certain values of the covariates for firm i observed up to time t, these conditional probabilities for observing rating category $r_{i,t}$ at time t are computed as: the joint probability of observing all available past ratings until time point $t-1$ and the current rating $r_{i,t}$, divided by the joint probability of all available past rating observations until time point $t-1$. Hence, knowing the joint distribution allows to predict ratings conditional on the previously observed ratings in the following way:

$$\mathbb{P}\left[Y_{i,t} = r_{i,t} \left| \left(\bigcap_{\substack{s<t,\\ s\in T_i}} Y_{i,s} = r_{i,s} \right) \right.\right] = \frac{\mathbb{P}\left(\bigcap_{s\in T_i} Y_{i,s} = r_{i,s}\right)}{\mathbb{P}\left(\bigcap_{\substack{s<t,\\ s\in T_i}} Y_{i,s} = r_{i,s}\right)}.$$

If the correlation of the errors along the time dimension is assumed to be zero i.e., $\rho = 0$, the conditional probability would clearly be equal to the marginal probability. This distinction among the conditional and marginal probabilities will be relevant in the evaluation of the longitudinal model. A better performance of the predictions based on conditional probabilities would point towards the fact that the past rating history is relevant for rating prediction. More discussion on this aspect can be found in Section 4.4.

3. Data and variable selection

3.1. Data

We collect S&P long-term issuer credit ratings from the Compustat–Capital IQ Credit Ratings database. In order to construct the firm-level variables we use the Compustat and CRSP databases together with the corresponding linking files. We compute an extensive list of financial ratios and market variables from the literature on corporate rating modeling (e.g., Alp, 2013; Baghai et al., 2014; Campbell et al., 2008; Tian et al., 2015) and consider further potential candidate ratios from *Financial Ratios Suite* tool of Wharton research data services (WRDS). The complete list of variables considered in this analysis is given in Table A.1.

We include in the analysis the universe of Compustat/CRSP US corporates which have at least one S&P rating observation in the period from 1985 to 2016. Financial, utility and real estate firms are excluded. The percentage of Compustat/CRSP firms with at least one S&P rating lies around 20% during the sample period. We merge the end-of-year ratings (available on December 31st) to the financial ratios on a calendar year basis, by assigning to the end-of-year ratings the latest Compustat entry (annual or quarterly) available before the end of the year. The market variables are computed using daily stock price data available between January 1st and December 31st of each year. All explanatory variables are winsorized at the 99th percentile; the ratios which can also take on negative values are winsorized at the 1st percentile.

In the merged sample, after eliminating the missing values in the covariates and in the ratings, we are left with 19 483 firm-year observations and 1800 firms. The panel of ratings is highly unbalanced, with many firms entering the sample after the beginning of the sample period or exiting the sample before 2016 due to rating withdrawal, default or other reasons such as mergers or acquisitions or unlisting. Moreover, some ratings are also temporarily missing due to temporary withdrawal or default. From all possible firm-year combinations, only 77.31% are observed. Only 50 firms have a rating in all sample years.

Fig. 1 shows the distribution of the S&P ratings over the sample period. S&P assigns corporates to 21 non-default rating classes where AAA, $AA+$, AA, $AA-$, $A+$, A, $A-$, $BBB+$, BBB, $BBB-$ are investment grade rating classes and $BB+$, BB, $BB-$, $CCC+$, CCC, $CCC-$, CC, C are the speculative grade classes. We merge in our sample the classes CCC/C due to low number of observations. It can be observed that the percentage of speculative grade firms increases around the end of the 90's and that the distribution becomes rather stable from 2006 onwards. The ratings change infrequently with 2515 downgrades (average number of downgrades per company is 1.397) and 1956 upgrades (average number of upgrades per company is 1.087) present in our sample.

3.2. Variable selection

Given that there is no consensus in the literature on which variables should be selected for inclusion in credit rating models, in this paper we employ a statistical approach for variable selection, namely the ordinal version of the LASSO. This method provides a well-established framework for selecting relevant variables in a univariate or static ordinal regression model. We perform the variable selection in a static rather than longitudinal model, due to the lack of available statistical software and for keeping computation time manageable, and perform a robustness exercise to check if the selected variables are justified in the longitudinal model. A brief discussion is provided at the end of Section 3.2.2.

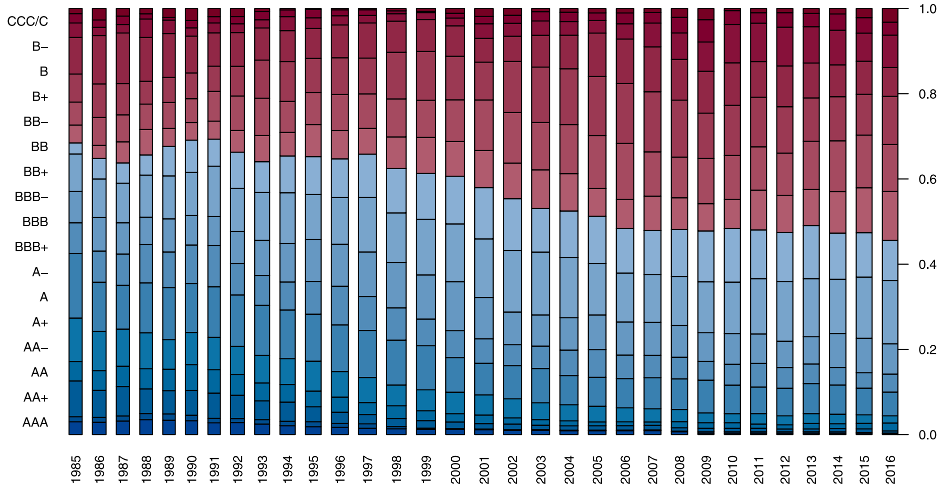

Fig. 1. This figure shows the rating distribution in the years 1985–2016. Blue (red) colors represent investment (speculative) grade ratings. The width of the bars is proportional to the number of observations in each year. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2.1. Ordinal LASSO

In order to select the most important variables from the extensive list in Table A.1, we employ the ordinal LASSO by building a static ordinal regression model to which a penalty term for the regression coefficients, namely the $\mathcal{L}_1$ norm, is added to the log-likelihood function. In the static model, the log-likelihood of the longitudinal model given in Eq. (2) is replaced by the log-likelihood of a univariate or static ordinal regression model given by the following sum over all firm-year observations in the sample:

$$\ell^{\text{iid}}(\boldsymbol{\theta}, \boldsymbol{\beta}) = \sum_{i=1}^{N} \sum_{t \in T_i} \log(\mathbb{P}(Y_{i,t} = r_{i,t} | \boldsymbol{x}_{i,t})) = \sum_{i=1}^{N} \sum_{t \in T_i} \log(\mathbb{P}(\widetilde{Y}_{i,t} \leq \theta_{r_{i,t}} | \boldsymbol{x}_{i,t}) - \mathbb{P}(\widetilde{Y}_{i,t} \leq \theta_{r_{i,t}-1}) | \boldsymbol{x}_{i,t}),$$

where for the logit link $\mathbb{P}(\widetilde{Y}_{i,t} \leq \theta_{r_{i,t}}) = \exp(\theta_{r_{i,t}} - \boldsymbol{x}_{i,t}^\top \boldsymbol{\beta})/(1 + \exp(\theta_{r_{i,t}} - \boldsymbol{x}_{i,t}^\top \boldsymbol{\beta}))$.

The ordinal LASSO imposes an $\mathcal{L}_1$ penalty on the regression coefficients $\boldsymbol{\beta}$ with the purpose of shrinkage and also variable selection. The negative log-likelihood of the model described above plus the penalty term $-\frac{1}{n}\ell^{\text{iid}}(\boldsymbol{\theta}, \boldsymbol{\beta}) + \lambda \sum_{j=1}^{p} |\beta_p|$ is minimized, where $\lambda \geq 0$ is a tuning parameter which controls the degree of shrinkage, with larger values introducing more shrinkage and hence smaller models. The value for this parameter can be chosen by AIC or BIC, or by cross-validation (if the model's ability to perform well on unseen data is of interest).

3.2.2. Results of the variable selection procedure

For a grid of values $\lambda_1, \lambda_2, \ldots, \lambda_M$ for the shrinkage parameter λ, the parameters of the ordinal LASSO are computed using the package **ordinalNet** (Wurm et al., 2017) for R (R Core Team, 2022). The results in this section are based on cross-validation adapted to the panel structure of the data. In using a cross-validation approach we are able to assess the ability of models containing different sets of variables to predict the ratings out-of-sample rather than in-sample. We set up the cross-validation exercise for time series (Bergmeir et al., 2018) by using ten years of data as training sample and the following year as a test sample. A total of 22 cross-validation samples are created using the rolling window approach (first training sample is 1985–1994). In our setting we are not only interested in the ability of the model to perform across time but also cross-sectionally i.e., across firms. Therefore, for each training sample we perform a repeated out-of-firm exercise, i.e., we randomly draw 10% of the firms in the training sample, remove all their observations from the training sample and add them to the test sample. We repeat this procedure 100 times. This out-of-sample analysis amounts to 2200 folds for each value of λ. The grid for the penalty parameter is chosen as suggested in Friedman et al. (2010).

As a criterion for choosing the optimal λ we employ weighted Cohen's κ_w (Cohen, 1968), a measure for the level of agreement among the predicted ratings and the observed ratings, with a value of zero indicating agreement by chance:

$$\kappa_w = 1 - \frac{\sum_{i=1}^{K} \sum_{j=1}^{K} w_{ij} o_{ij}}{\sum_{i=1}^{K} \sum_{j=1}^{K} w_{ij} e_{ij}},$$

where w_{ij}, o_{ij} and e_{ij} denote the weights, observed and expected elements of the contingency matrix among predicted ratings and observed ratings. Each of the folds serves once as a test sample and κ_w is computed for each fold for the grid of λ values. The weights w_{ii} corresponding to the diagonal are set to zero, while the off-diagonal weights indicate the "seriousness" of the disagreement. For an ordinal scale, squared weights $w_{ij} = 1 - (i - j)^2/(K - 1)^2$ i.e., squared distances from perfect agreement, are typically employed to more strongly penalize values far off from the diagonal.

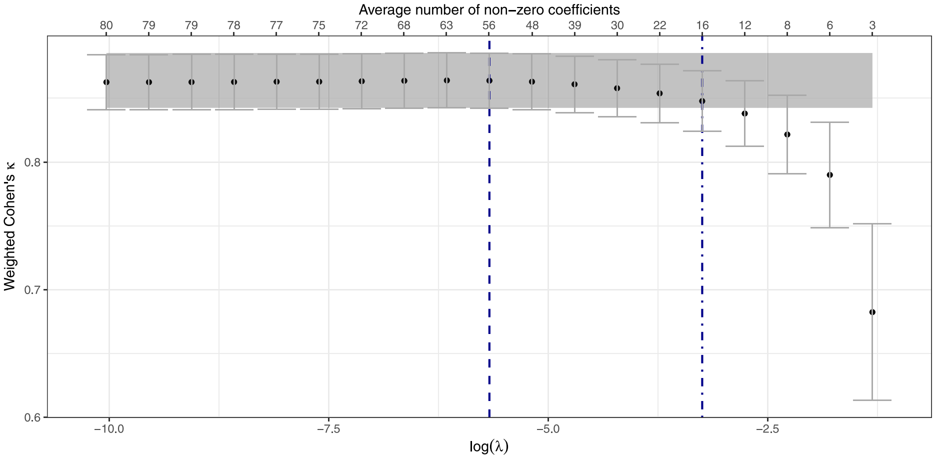

Fig. 2. This figure illustrates the mean and the 2.5% and 97.5% quantile of κ_w over the 2200 folds for different values of the shrinkage parameter λ. The dashed line indicates the (log) value of λ for which the agreement among predicted and the observed ratings is maximal and the gray band indicates the 95% cross-validated interval of the maximum κ_w. The dash-dotted line corresponds to the *one-standard-error* rule.

Fig. 2 shows the mean and the 95% cross-validated interval of κ_w with squared weights for given λ values. The maximal average κ_w is achieved for a model containing an average of 56 variables and is therefore not appropriate for the purpose of variable selection as the model size is still prohibitively large. Another alternative in the literature is the "one-standard-error" heuristic (Hastie et al., 2009), where the optimal model is obtained by choosing the largest value λ_m (i.e., smallest model) such that the average κ_w is within one standard deviation from the maximum. As can be seen from Fig. 2, the model chosen by the "one-standard-error" rule contains on average 16 variables across the 2200 folds and the value of the penalty term is in this case $\lambda^{1se} = 0.039$. For each value of λ we can also investigate how often each variable has been selected in the 2200 folds. These inclusion probabilities are illustrated in Fig. 3. The cross-validation exercise thus allows us to check the relevance of the variables for the rating prediction task by looking at which stage (i.e., for which λ and implicitly model size) each variable enters the model. For example, it can be seen that the relative size of the company and retained earnings to assets are the most important variables, included in all models for all values of λ. On the other hand the capitalization ratio is included in few folds regardless of the shrinkage parameter. Similarly, ratios like the current ratio or operating profit margin after depreciation are only included in large models which contain the vast majority of the variables.

In defining the final variables to enter the model, we choose the tuning parameter $\lambda = \lambda^{1se}$. For this value of λ, 4 variables are included in all 2200 folds (*div_payer*, *re_at*, *SIGMA*, *RSIZE*), 32 variables were never included and 67 are included in less than 50% of the folds. We propose that the final model includes the 13 variables which are included in at least 50% of the folds: *BETA*, *che_mta*, *convdebt_at*, *debt_ebitda*, *div_payer*, *dltt_at*, *ebit_at*, *equity_invcap*, *lt_mta*, *NYSE_SIZE_PERC*, *re_at*, *RSIZE*, and *SIGMA*. Table 1 contains the descriptive statistics for these ratios for the rating classes aggregated into investment versus speculative grade as well as for the entire sample. Table A.2 in Appendix shows the average values for each of the sample years. It is to be noted that the statistical variable selection exercise delivers ratios which are in accordance to the literature to a large extent. We observe that the ratios selected measure different aspects of firm performance such as size (*RSIZE*, *NYSE_SIZE_PERC*), profitability (*ebit_at*, *re_at*), liquidity (*che_mta*), leverage and capital structure (*convdebt_at*, *debt_ebitda*, *dltt_at*, *equity_invcap*, *lt_mta*), risk based on market prices (*BETA*, *SIGMA*) and whether the company is a dividend payer (*div_payer*). Eight of these ratios are also employed in Alp (2013). In our exercise, the group of leverage and capital structure ratios is larger compared to other models in the literature (e.g., Alp, 2013 includes only long-term and total debt leverage ratios in their model). Additionally, variables measuring cash-flow, tangibility and interest rate coverage are not present in the final list of 13 ratios.

In order to assess the robustness of the variable selection we performed several exercises. We employed the percentage of correct predictions as a measure of out-of-sample performance instead of weighted Cohen's κ_w when choosing the optimal shrinkage parameter but this analysis delivered no relevant differences in the selected variables. Moreover, to justify the choice of λ^{1se}, we estimate for each value of λ the longitudinal model containing the variables included in more than 50% of the folds and compare the in-sample information criteria CLAIC and CLBIC. The results reveal that, while larger models do perform better in-sample, the improvement in the information criteria is marginal for models larger than the one corresponding to $\lambda^{1se} = 0.039$. This exercise serves as a robustness check for investigating whether the variables selected by the LASSO in a static framework are justifiable in the longitudinal model.

4. Empirical results

Using the variables selected by the LASSO, we estimate and present the results of the longitudinal regression model. Additionally, we compare this model with a static version in terms of in-sample and out-of-sample performance. We also estimate longitudinal

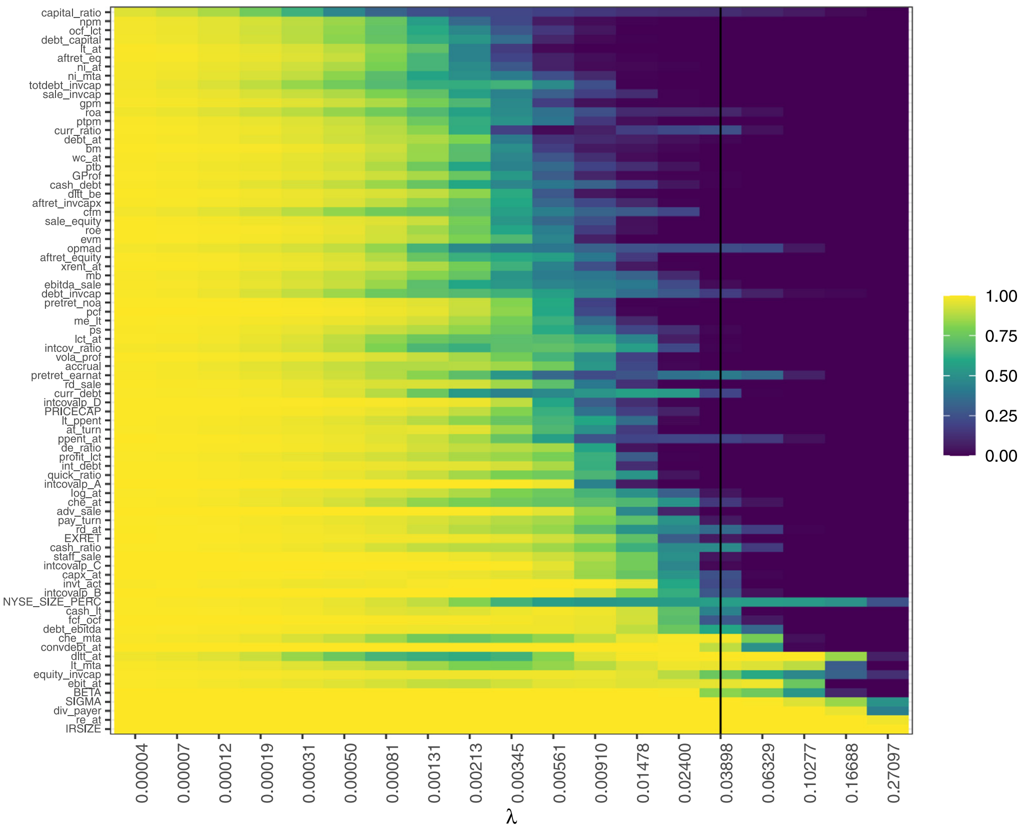

Fig. 3. This figure illustrates the proportion out of the 2200 folds in which the variables have been selected for inclusion for the λ grid values. The vertical line indicates $\lambda^{1se} = 0.039$.

and static versions of models containing the ratios in Alp (2013) and Baghai et al. (2014) to be able to assess the value added by the variable selection exercise in comparison to these alternative models. Finally, we extend the framework by allowing for time-varying regression coefficients and discuss the implications of such a model.

4.1. Model fit

We proceed with the analysis by fitting a multivariate ordinal logistic regression model with $AR(1)$ errors over the whole sample period. We keep the threshold coefficients fixed for all the time points. In general, the coefficients of the variables have the expected sign (see Table A.3). We observe that higher levels of cash holdings to assets *che_mta*, higher debt- and liability-related ratios (as measured by convertible debt to assets *convdebt_at*, *debt_ebitda*, long-term debt as percentage of total assets *dltt_at*, liabilities to market assets *lt_mta*) as well as higher systematic risk and idiosyncratic risk are associated with lower ratings. Moreover, large firms, dividend payers, firms which exhibit large profitability ratios (as measured by EBIT to assets *ebit_at* and retained earnings to assets *re_at* ratios), as well as higher equity ratios as given by equity to invested capital (*equity_invcap*) receive on average better ratings. Finally, for the autocorrelation parameter ρ of the error structure, we obtain a value of 0.92 (with a standard error of 0.01), which is strong evidence of serial correlation in the credit ratings.

We also employed a multivariate normal distribution for the errors in Eq. (1) instead of the multivariate logistic distribution. As goodness-of-fit measures we compare the CLAIC, CLBIC and McFadden's adjusted pseudo R^2 (McFadden, 1977). Note that the value of this measure is substantially lower compared to an ordinary R^2, with values of 0.2 to 0.4 indicating an excellent model fit. We find that the multivariate logit link is slightly favored by all these measures: a CLAIC of 1 158 911.55 for the logit link is lower compared to 1 190 723.39 for the probit link, as well as a lower CLBIC for the logit link (1 170 103.39) compared to the probit link (1 209 155.13); the pseudo R^2 decreases slightly from 0.2873 to 0.2692 when employing the probit link.

Table 2 contains all in-sample measures of fit for the longitudinal and static models containing the variables selected by the LASSO, the variables proposed in Alp (2013) and the ones in Baghai et al. (2014). The value added by the longitudinal specification is reflected in the improvement of the goodness-of-fit measures when comparing the longitudinal models to the static versions. The longitudinal model with the LASSO variables performs the best in-sample in terms of CLAIC, CLBIC and pseudo R^2, followed closely by the longitudinal model with the variables in Alp (2013). The model containing the ratios in Baghai et al. (2014) performs the poorest, but it is to be mentioned that the original paper employed a static linear regression model for the purpose of rating

Table 1
Descriptive statistics of the selected financial ratios for the firms for investment grade (IG) and speculative grade (SG) ratings as well as for the entire sample (Full).

Ratio	Rating	Mean	1st Qu.	Median	3rd Qu.	Std. dev.
BETA	SG	1.1434	0.7324	1.1084	1.5077	0.5560
	IG	0.9297	0.6666	0.9165	1.1698	0.3802
	Full	1.0215	0.6901	0.9824	1.2988	0.4759
che_mta	SG	0.0618	0.0150	0.0370	0.0830	0.0736
	IG	0.0399	0.0110	0.0250	0.0520	0.0495
	Full	0.0493	0.0120	0.0290	0.0630	0.0620
convdebt_at	SG	0.0378	0.0000	0.0000	0.0350	0.0791
	IG	0.0111	0.0000	0.0000	0.0000	0.0373
	Full	0.0227	0.0000	0.0000	0.0000	0.0606
debt_ebitda	SG	3.9281	2.0060	3.3420	5.1110	6.2090
	IG	2.0141	1.0760	1.6940	2.5205	2.2532
	Full	2.8356	1.2730	2.1680	3.6590	4.5101
div_payer	SG	0.3611	0.0000	0.0000	1.0000	0.4803
	IG	0.8655	1.0000	1.0000	1.0000	0.3412
	Full	0.6489	0.0000	1.0000	1.0000	0.4773
dltt_at	SG	0.3767	0.2490	0.3590	0.4830	0.1808
	IG	0.2247	0.1390	0.2140	0.2970	0.1221
	Full	0.2900	0.1710	0.2640	0.3800	0.1679
ebit_at	SG	0.0674	0.0350	0.0710	0.1070	0.0896
	IG	0.1152	0.0740	0.1070	0.1490	0.0654
	Full	0.0946	0.0560	0.0920	0.1330	0.0803
equity_invcap	SG	0.4102	0.2800	0.4510	0.5990	0.2903
	IG	0.6211	0.5150	0.6380	0.7542	0.2039
	Full	0.5307	0.4060	0.5690	0.7060	0.2660
log(RSIZE)	SG	−9.6603	−10.3863	−9.6342	−8.9141	1.1490
	IG	−7.6708	−8.5779	−7.7236	−6.8054	1.2548
	Full	−8.5261	−9.6044	−8.5858	−7.4410	1.5605
lt_mta	SG	0.5594	0.4090	0.5610	0.7120	0.2020
	IG	0.3912	0.2600	0.3780	0.5060	0.1718
	Full	0.4635	0.3080	0.4470	0.6048	0.2033
NYSE_SIZE_PERC	SG	0.5123	0.3538	0.5264	0.6833	0.2233
	IG	0.8198	0.7378	0.8554	0.9382	0.1500
	Full	0.6876	0.5342	0.7392	0.8861	0.2397
re_at	SG	−0.0013	−0.0762	0.0660	0.1980	0.4528
	IG	0.3135	0.1710	0.3030	0.4420	0.2257
	Full	0.1779	0.0505	0.2030	0.3640	0.3763
SIGMA	SG	0.0277	0.0186	0.0246	0.0328	0.0132
	IG	0.0168	0.0122	0.0155	0.0199	0.0066
	Full	0.0215	0.0139	0.0186	0.0258	0.0113

prediction, rather than an ordinal model. We also compare the in-sample model performance in terms of prediction accuracy. The importance of modeling the autocorrelation is also reflected when comparing the marginal predictions of the proposed model to the conditional predictions. We generate predictions based on the conditional probabilities of observing each rating given the past rating observations of the firm (as described in Section 2) as well as the predictions based on unconditional (i.e., marginal) probabilities. The predicted rating grade in each case is the one with the largest predicted probability. Given that univariate marginal predictions do not take the multivariate structure into account, we can observe that the weighted Cohen's κ_w is markedly higher for the conditional predictions. The predictions of the static model are similar in magnitude to the marginal predictions of the longitudinal model. For the predictions within $k = 0, 1, 2$ rating notches we find major differences between the conditional versus marginal approaches and observe a significant out-performance of the conditional predictions. For $k \geq 3$, the differences get smaller and we find that within $k = 6$ notches all models achieve an accuracy of above 95%.

Finally, for all employed measures we see that the relative importance of adding the longitudinal structure in the model specification is larger than the improvement gained by the variable selection exercise over the variables employed in Alp (2013) (e.g., the reduction in CLAIC when moving from the static to the dynamic specification is 5.53%, while the reduction in CLAIC when moving from the model with the variables in Alp (2013) to the model with the LASSO selected variables is 1.05%).

4.2. Out-of-sample performance analysis

In order to investigate the out-of-sample model performance, we perform a rolling windows prediction exercise where we iteratively fit the model on a ten years window and predict the ratings of the following year. We benchmark the proposed longitudinal model against the static specification. We measure the out-of-sample prediction performance by κ_w and the percentage of correctly predicted categories within $k = 0, 1, 2, 3$ rating notches. Table 3 presents the results averaged over the whole test period while results on a yearly basis are presented in Fig. 4. When comparing the conditional predictions to the predictions of the static model, we

Table 2
This table presents in-sample model performance such as CLAIC, CLBIC, adjusted pseudo R^2, weighted Cohen's κ_w and the percentage of correct predictions within $k = 0, 1, 2, 3, 6$ rating notches for conditional and marginal predictions for both the longitudinal model and the static model (i.e., without autocorrelation) with logit link. For both the static and longitudinal models we report results when using the ratios selected by the LASSO, the ratios in Alp (2013) and the ratios in Baghai et al. (2014). Note that Alp (2013) as well as Baghai et al. (2014) employ the static version in their work. Also, for the static model, both conditional and marginal predictions correspond to the predictions obtained from a univariate model.

Measure	LASSO		Alp (2013)		Baghai et al. (2014)	
	Longitudinal	Static	Longitudinal	Static	Longitudinal	Static
CLAIC	1 158 911.55	1 226 749.52	1 171 257.7	1 238 537.72	1 492 114.61	1 762 871.46
CLBIC	1 170 103.39	1 241 073.47	1 183 541.83	1 253 772.14	1 506 359.15	1 779 632.1
Adj. pseudo R^2	0.2873	0.2461	0.2799	0.239	0.0824	−0.0841
Conditional predictions						
κ_w	0.9667	0.8315	0.9661	0.8189	0.8829	0.5805
% of correct predictions	55.64%	25.30%	55.24%	24.68%	31.79%	17.83%
within $k = 1$ notches	92.80%	58.39%	92.48%	56.70%	69.11%	42.96%
within $k = 2$ notches	98.48%	78.23%	98.42%	76.59%	87.24%	59.55%
within $k = 3$ notches	99.61%	91.60%	99.65%	90.28%	96.56%	76.53%
within $k = 6$ notches	99.97%	99.63%	99.96%	99.36%	99.74%	96.83%
Marginal predictions						
κ_w	0.8231	0.8315	0.8075	0.8189	0.6563	0.5805
% of correct predictions	25.13%	25.30%	24.18%	24.68%	18.63%	17.83%
within $k = 1$ notches	57.61%	58.39%	55.81%	56.70%	45.48%	42.96%
within $k = 2$ notches	77.49%	78.23%	75.30%	76.59%	63.05%	59.55%
within $k = 3$ notches	91.21%	91.60%	89.93%	90.28%	80.17%	76.53%
within $k = 6$ notches	99.63%	99.63%	99.45%	99.36%	98.07%	96.83%

Table 3
This table presents out-of-sample weighted Cohen's κ_w and the percentage of correct predictions within $k = 0, 1, 2, 3, 6$ rating notches for conditional and marginal predictions for both the longitudinal model and the static model (i.e., without autocorrelation) with logit link for three sets of variables: the ratios selected by the LASSO, the ratios in Alp (2013) and the ratios in Baghai et al. (2014). Note that for the static model, both conditional and marginal predictions correspond to the predictions obtained from a univariate model.

Measure	LASSO		Alp (2013)		Baghai et al. (2014)	
	Longitudinal	Static	Longitudinal	Static	Longitudinal	Static
Conditional predictions						
κ_w	0.9574	0.8438	0.9553	0.8324	0.3834	0.0618
% of correct predictions	56.29%	27.14%	55.60%	25.96%	14.00%	3.48%
within $k = 1$ notches	91.78%	63.39%	91.28%	61.77%	30.48%	7.65%
within $k = 2$ notches	97.70%	83.97%	97.39%	82.78%	41.76%	11.31%
within $k = 3$ notches	99.29%	94.78%	99.26%	94.30%	51.79%	15.59%
within $k = 6$ notches	100.00%	99.88%	99.98%	99.80%	72.92%	35.43%
Marginal predictions						
κ_w	0.8326	0.8438	0.8198	0.8324	0.0586	0.0618
% of correct predictions	25.96%	27.14%	25.69%	25.96%	2.63%	3.48%
within $k = 1$ notches	60.53%	63.39%	59.89%	61.77%	5.33%	7.65%
within $k = 2$ notches	81.63%	83.97%	80.45%	82.78%	8.41%	11.31%
within $k = 3$ notches	93.85%	94.78%	93.18%	94.30%	12.06%	15.59%
within $k = 6$ notches	99.88%	99.88%	99.78%	99.80%	31.17%	35.43%

find that, as in the in-sample analysis, conditional predictions are superior to ones from the static model for all the years and for all three variable sets. Moreover, the out-of-sample performance of the longitudinal model with LASSO variables increases slightly after the financial crisis. A similar pattern is observed for the corresponding model with the variables in Alp (2013). As in the in-sample evaluation, the relative importance of adding the longitudinal structure in the model specification is larger than the improvement gained by the variable selection exercise over the variables employed in Alp (2013).

4.3. Model extension: including time-varying coefficients

The longitudinal model can be easily extended to allow for time-varying regression coefficients. Analyzing the stability of the coefficients over time in a credit rating model can be relevant in order to understand whether there has been a change in the rating behavior of S&P. We fit a multivariate ordinal regression model with $AR(1)$ errors with the following specification for the process on the latent scale:

$$\widetilde{Y}_{i,t} = \widetilde{\boldsymbol{x}}_{i,t}^{\top}\boldsymbol{\beta}_t + \epsilon_{i,t}. \quad (3)$$

The rest of the assumptions remain the same as in the proposed model in Eq. (1).

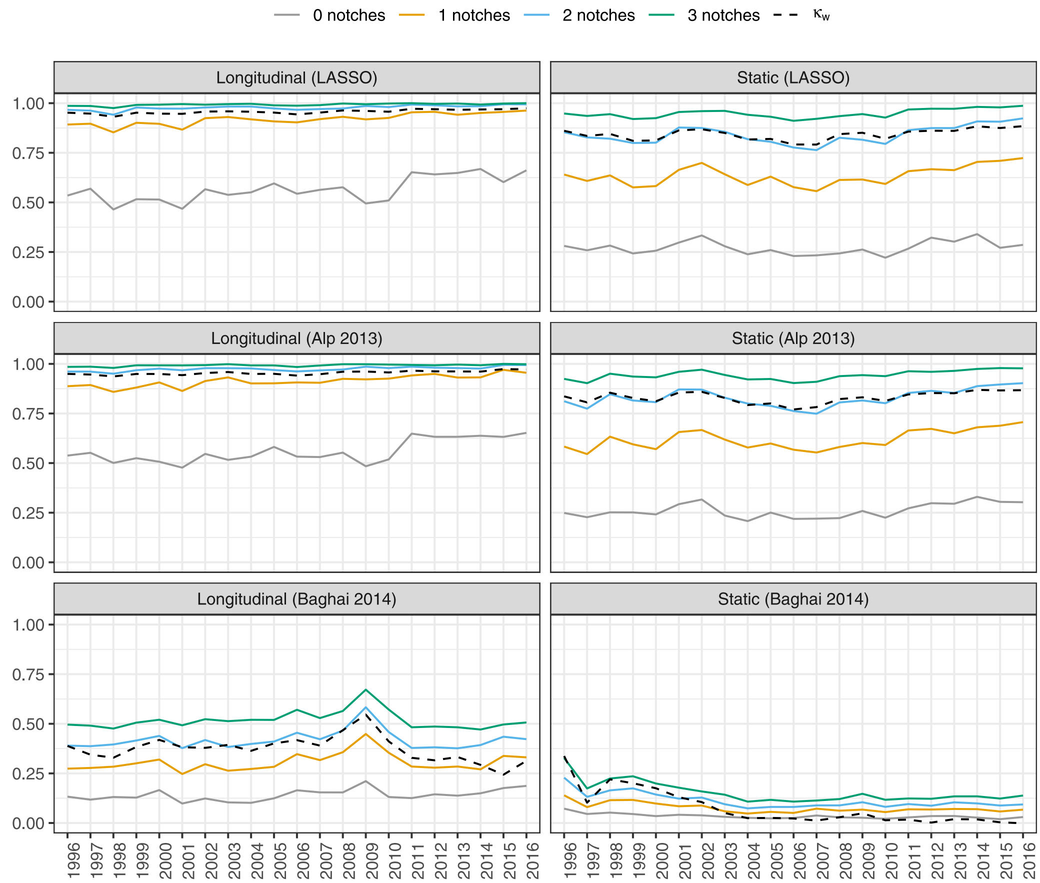

Fig. 4. This figure shows the out-of-sample performance measures where 10 years of data are used as training data in a rolling window fashion and the years 1995–2016 serve as one-year ahead test samples. The measures illustrated are weighted Cohen's κ_w (black dashed line) and accuracy (i.e., proportion of correct predictions) within $k = 0, 1, 2, 3$ rating notches. The panels correspond to the conditional predictions from the longitudinal model and to the predictions from the static model, separated for the three sets of ratios.

Fig. 5 presents the scaled coefficients which can be interpreted as the number of rating notches that a firm would deteriorate its rating given a one-standard-deviation increase in the explanatory variable. The FRED recession indicator series was used for identifying crisis periods 1990–1991, 2001–2002, 2007–2009 which are also marked in the figure. We observe that coefficients for the size ratios *NYSE_SIZE_PERC* and log(*RSIZE*) are mostly stable and fluctuate around the mean, with a slight increase for log(*RSIZE*) until 1995. For retained earnings to assets ratio *re_at* the coefficient exhibits a drop during the 1990—1991 recession and another drop in 2007. As in the case of cash holdings, it has been documented that after the financial crisis of 2008, North-American companies, on average, showed an abnormally high level of earning retention (Paulo, 2018), which can partially explain the behavior. The decrease in absolute value until 2008 and the rather constant evolution of the long-term debt to assets ratio *dltt_at* thereafter can be seen as a shift in the focus of S&P towards equity rather than long-term debt in their leverage assessment. This shift could be caused by the unprecedented monetary expansion following the financial crisis, which increased debt levels throughout all sectors, together with the increased use of short-term debt.[4] The coefficient of idiosyncratic risk *SIGMA* exhibits a decrease in absolute value (and hence an increased impact) before the crisis. The ratio *convdebt_at* increased in absolute value after 2010 while *debt_ebitda* seems to be a good predictor for the ratings after the dot-com bubble and up to the financial crisis 2007–2009 but less so in the other sample years.

As in the model with equal coefficients, for the autocorrelation value ρ we obtain a value of 0.92 (standard error 0.03). When comparing the model with equal coefficients with the time-varying coefficients model we find that the goodness-of-fit measures slightly prefer the time-varying coefficients model. We observe a smaller CLAIC (from 1 158 911.55 in the equal coefficients model to 1 081 129.63 in the time-varying model) as well as a smaller CLBIC (from 1 170 103.39 to 1 137 869.00) and an increase of the adjusted pseudo R^2 (from 0.29 to 0.34). In terms of in-sample predictive measures, models perform similarly: weighted Cohen's κ_w is 0.9707 (time-varying) and 0.9667 (equal) and the percentage of correctly predicted categories is 0.5865 (time-varying) and

[4] See https://regulationbodyofknowledge.org/wp-content/uploads/2013/03/StandardAndPoors_Corporate_Ratings_Criteria.pdf S&P's corporate ratings criteria.

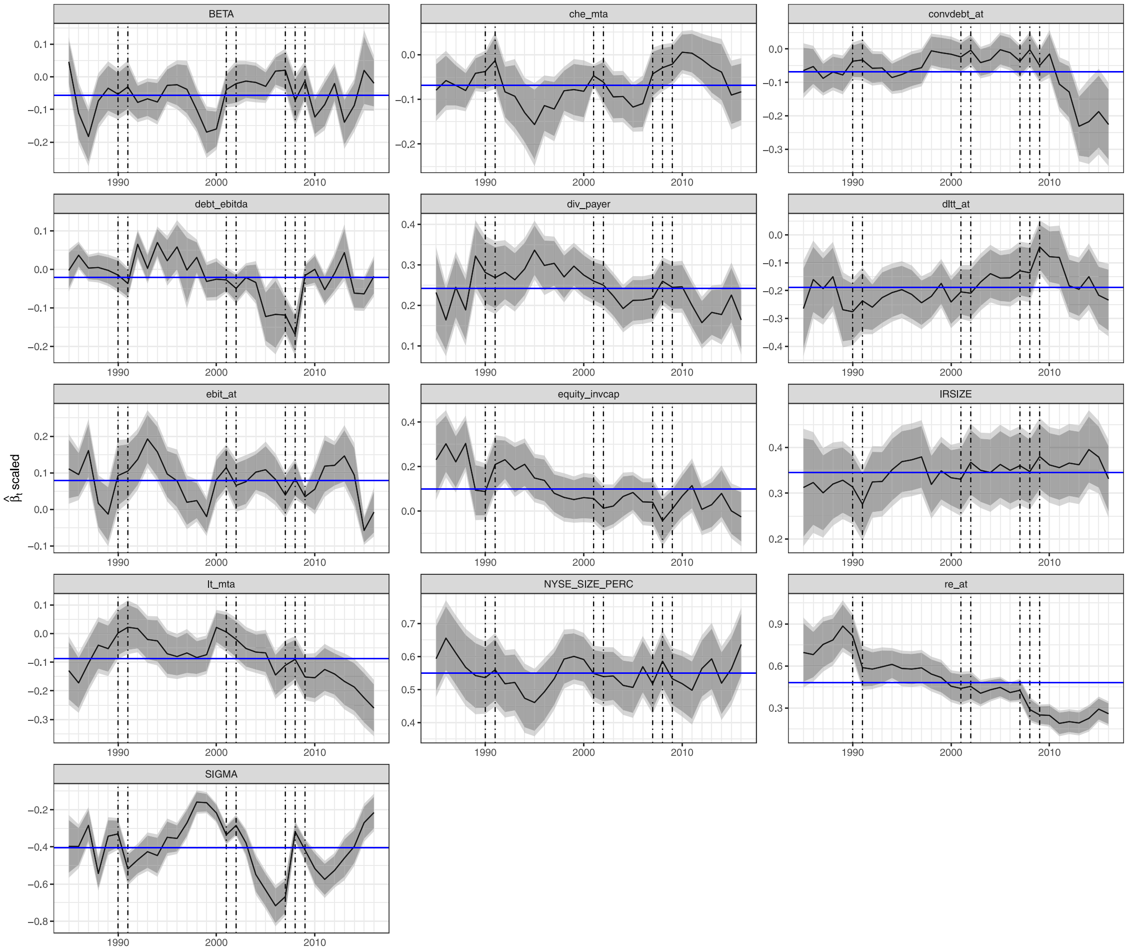

Fig. 5. This figure shows the time-series of the estimated regression coefficients together with average value (solid blue line) and the corresponding 95% (dark gray) and 90% (light gray) confidence bands. For comparison purposes, the estimated coefficients have been scaled by multiplying the estimated $\hat{\beta}_{pt}$ with the standard deviation of the pth variable and dividing by the rating notch length $|\theta_{CC} - \theta_{AAA}|/K$. The vertical dash-dotted lines indicate the crisis periods 1990–1991, 2001–2002, 2007–2009. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

0.5564 (equal). Note however that the time-varying coefficients model serves an inference purpose, where the goal is understanding the changes in the underlying relationships rather than achieving a high predictive performance. For this reason, we omit an out-of-sample analysis. Moreover, for prediction purposes one would need to further assume some dynamics on the time-varying coefficients in order to expect an improvement of the out-of-sample performance.

4.4. Further discussion of results

The superior performance of the longitudinal model, especially in terms of out-of-sample performance based on conditional predictions, implies that the probability of observing a rating in a given period is influenced by the past rating history. This is the case either because the previous ratings cause a change in the latent creditworthiness (e.g., good ratings decrease the costs of funding and hence lead to better credit quality Manso, 2013) and/or because of heterogeneity caused by unobserved factors which makes the ratings serially correlated. To investigate this aspect for our longitudinal model, we additionally estimate the model specification with equal coefficients but add ratings lagged by one year as covariates. Results can be found in Table A.4 in Appendix. This model does not improve model performance in terms of conditional predictions (see Table A.5 for a comparison of the in-sample measures based on conditional and marginal predictions), but we do observe significant effects for the rating dummies and a large persistence parameter of $\rho = 0.77$. This shows that heterogeneity as well as past rating dependence contribute to the superior performance of the longitudinal model. The high persistence as captured by ρ can be indicative of a high reliance of the rating agencies on soft facts in the rating process.

Finally, we investigate how the modeling framework can be used for identifying characteristics of the rating behavior such as procyclicality. For this purpose we look at the out-of-sample performance of various models and regressor specifications, as in the analysis presented in Section 4.2. More specifically, we investigate whether the actual ratings exceed the model-predicted ratings

Table 4
This table presents the percentage of observations where the (i) assigned ratings Y are worse than the predicted ratings $\hat{Y}$ ($\mathbb{P}(Y < \hat{Y})$), (ii) assigned ratings are equal to the predicted ratings ($\mathbb{P}(Y = \hat{Y})$), (iii) assigned ratings are higher than the predicted ratings ($\mathbb{P}(Y > \hat{Y})$), for a variety of models in the out-of-sample exercise.

Model	Crisis periods			Non-crisis periods		
	$\mathbb{P}(Y < \hat{Y})$	$\mathbb{P}(Y = \hat{Y})$	$\mathbb{P}(Y > \hat{Y})$	$\mathbb{P}(Y < \hat{Y})$	$\mathbb{P}(Y = \hat{Y})$	$\mathbb{P}(Y > \hat{Y})$
Longitudinal (LASSO)	30.95%	53.36%	15.68%	25.36%	57.21%	17.09%
Static (LASSO)	45.82%	27.52%	26.79%	42.81%	27.03%	30.08%
Longitudinal (Alp, 2013)	31.60%	51.82%	16.59%	25.79%	56.78%	17.10%
Static (Alp, 2013)	47.31%	26.35%	26.47%	45.26%	25.84%	28.74%
Longitudinal (Baghai et al., 2014)	70.54%	14.63%	14.64%	73.64%	13.80%	12.48%
Static (Baghai et al., 2014)	91.01%	3.49%	5.53%	90.50%	3.48%	6.05%

in non-crisis periods and whether during the crisis the assigned ratings are lower than the predicted ratings ($Y < \hat{Y}$). Such a pattern would be indicative of a procyclical rating behavior, with rating agencies being too fast to downgrade in a recession. We report in Table 4 the fraction of predicted ratings which are equal to the observed ratings, the percentage of ratings overestimated and the percentage of ratings underestimated by different models for the crisis and non-crisis period. The percentage of correct predictions is substantially higher for the longitudinal models and is highest for the longitudinal LASSO model proposed in this paper (as was also shown before in Table 3). The results show that the longitudinal specification for the different variable groups is better in both crisis and non-crisis periods (also to be seen in Table 3). When looking at the percentage of over- and underestimations we see that in crisis times there is an increase in terms of overestimations. The probability of the observed rating being worse than the predicted rating increases from 25% in the non-crisis periods to 31% in crisis times for the longitudinal LASSO model. This suggests that the rating agencies are not fully through the cycle and show to some extent aspects of procyclical behavior. Note that in the static models, the differences between crisis versus non-crisis periods are less marked, but these models do not exhibit a competitive predictive out-of-sample performance.

5. Conclusion

In this paper we propose a longitudinal model for predicting ratings on an ordinal scale. By modeling the multivariate distribution over all years, we are able to assess predictions conditional on the previously observed ratings and find an improvement in the predictive performance in-sample and out-of-sample.

As a first step of the analysis, we employ the ordinal LASSO to select the relevant variables out of a comprehensive list of 78 variables. The analysis suggests the inclusion of 13 variables into the ordinal models, namely: systematic and idiosyncratic risk, cash and cash equivalents to market assets, debt to EBITDA ratio, a flag indicating whether the firm is a dividend payer, earnings before interest and taxes divided by assets, equity as percentage of invested capital, liabilities to market assets ratio, long-term debt as percentage of total assets, convertible debt to assets ratio, two size ratios and retained earnings to assets ratio. The ordinal LASSO procedure also provides guidance into the relevance of the variables for this type of models, as the variables can be ranked in terms of their model inclusion probabilities for a given model size. Moreover, our analysis shows that the out-of-sample predictive performance of the ordinal model does not significantly decrease for more parsimonious models.

After identifying the relevant determinants of corporate credit ratings, we set up the longitudinal ordinal logit model with autoregressive errors and perform an extensive in and out-of-sample performance evaluation. We find that accounting for the longitudinal structure increases all the goodness-of-fit as well as prediction performance measures. In our application, the multivariate logit link exhibits a better fit than more common multivariate probit link while allowing for a meaningful interpretation of the coefficients in terms of log-odds ratios.

Additionally, we extend the model by allowing for time-varying coefficients and investigate how the relationship between the selected variables and the S&P corporate credit ratings changes over time. We find that for variables such as retained earnings to assets ratio the effect on the ratings diminishes over the sample period, while the convertible debt ratio and liability to market assets ratio increases their impact on the corporate credit ratings of S&P after the financial crisis 2007–2009. On the other hand, variables such as debt to EBITDA or idiosyncratic risk have a stronger impact in the period between the dot-com bubble and the start of the 2007–2009 financial crisis. Variables related to size have a large rather constant effect on the ratings. These findings suggest that S&P's assessment of risk has changed after the financial crisis, given also the change in the market conditions. Furthermore, we observe that the longitudinal model with LASSO selected variables tends to overestimate the ratings in crisis periods more than in non-crisis periods, which is an indication of procyclical aspects in the rating behavior of S&P.

Future research includes extending the analysis to also incorporate default information as well as a multi-dimensional analysis of the time variation of the coefficients for several credit rating agencies including e.g., Moody's and Fitch while accounting for the inter-dependencies among these raters.

CRediT authorship contribution statement

Rainer Hirk: Conceptualization, Methodology, Software, Writing. **Laura Vana:** Conceptualization, Methodology, Software, Writing. **Kurt Hornik:** Supervision.

Table A.1
This table displays the ratios used in the analysis.

Code	Ratio	Description
Capitalization ratios		
capital_ratio	Capitalization ratio	Long-term debt (`dltt`) divided by the sum of long-term debt (`dltt`), common equity (`ceq`) and preferred stock (`pstk`)
debt_invcap	Long-term debt as % of invested capital	Long-term debt `dltt` divided by `icapt`
equity_invcap	Common equity as % of invested capital	Common equity `ceq` divided by invested capital (`icapt`)
totdebt_invcap	Total debt as % of invested capital	Long-term debt (`dltt`) plus short-term debt (`dlc`) divided by `icapt`
Financial soundness ratios		
cash_lt	Cash balance as % of liabilities	Cash and cash equivalents (`che`) divided by total liabilities (`lt`)
convdebt_at[c]	Convertible debt ratio	Convertible debt `dcvt` divided by total assets `at`
debt_at[bcd]	Debt ratio	Total debt (`dltt+dlc`) divided by total assets `at`
debt_ebitda[c]	Total debt to EBITDA	Total debt `dltt+dlc` divided by earnings before interest, taxes, depreciation and amortization `ebitda`
fcf_ocf	Free cash flow to operating cash flow	Operating cash flow `ocf` minus capital expenditures `capx` divided by `ocf`
int_debt	Interest paid on long-term debt	Interest expenses (`xint`) divided by average long-term debt (`dltt`) in the current and previous year
invt_act	Inventory to current assets	Inventories `invt` divided by current assets `act`
lct_at[d]	Current liabilities to assets	Current liabilities `lct` divided by total assets `at`
lt_ppent	Liabilities relative to tangible assets	Total liabilities `lt` divided by net property, plant and equipment `ppent`
ocf_lct	Operating cash flow to current liabilities	Operating cash flow `ocf` divided by total current liabilities `lct`
ocf_lt	Operating cash flow to total liabilities	Operating cash flow `ocf` divided by total liabilities `lt`
profit_lct	Operating profit to current liabilities	Operating income before depreciation `oibdp` divided by total current liabilities `lct`
Solvency ratios		
curr_debt	Current liabilities as % of total liabilities	Current liabilities `lct` divided by total liabilities `lt`
de_ratio	Debt to equity ratio	Total liabilities `lt` divided by the sum of common equity `ceq` and preferred stock `pstk`
debt_capital	Debt to capital ratio	Total debt `dltt+dlc` divided by the sum of accounts payable `ap`, total debt `dltt+dlc`, common equity `ceq` and preferred stock `pstk`
dltt_at[b]	Long-term debt to assets ratio	Long-term debt `dltt` divided by total assets `at`
dltt_be	Long-term debt to equity ratio	Long-term debt `dltt` divided by book equity `seq`
intcov_ratio[c]	Interest coverage ratio	Earnings before interest and taxes `ebit` divided by `xint`
intcovalp[b]	Interest coverage ratio (2)	Interest expenses `xint` plus operating income after depreciation `oiadp` divided by interest expenses `xint`; this ratio is then split into four continuous variables: A contains values between [0,5], B: (5,10], C: (10,20], D: (20,100].
lt_at	Liabilities to assets ratio	Total liabilities `lt` to total assets `at`
lt_mta[a]	Liabilities to market assets	Total liabilities `lt` divided by total market assets (`lt+prcc*csho`)
me_lt	Market equity to total liabilities	Market equity (closing share price `prcc` times common shares outstanding `csho`) divided by total liabilities `lt`
Liquidity ratios		
cash_ratio	Cash ratio	Cash and cash equivalents `che` divided by total current liabilities `lct`
che_at[bc]	Cash to assets ratio	Cash and cash equivalents `che` divided by total assets `at`
che_mta[a]	Cash to market assets ratio	Cash and cash equivalents `che` divided by total market assets `lt + prcc x csho`
curr_ratio	Current ratio	Current assets `act` divided by current liabilities `lct`
quick_ratio	Quick ratio	Current assets `act` minus inventories `invt` divided by current liabilities `lct`
wc_at	Working capital to assets	Working capital `act-lct` divided by total assets `at`
Profitability ratios and rates of return		
aftret_eq	After-tax return on average common equity	Income before extraordinary items (available for common) `ibcom` divided by average common equity `ceq` in the current and previous year.

(*continued on next page*)

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors acknowledge funding from the OeNB Jubiläumsfonds for the project no. 18482 "Multivariate ordinal regression models for enhanced credit risk modeling".

Appendix. Tables

See Tables A.1–A.5.

Table A.1 (*continued*).

Code	Ratio	Description
Capitalization ratios		
aftret_equity	After-tax return on total stockholders equity	Income before extraordinary items `ib` divided by average book equity `seq` in the current and previous year.
aftret_invcapx	After-tax return on invested capital	Income before extraordinary items `ib` plus interest expenses `xint` divided by average invested capital `icapt` in the current and previous year.
cfm	Cash flow margin	Cash flow income before extraordinary items `ibc` plus cash flow depreciation `dpc` divided by total sales `sale`
ebit_at	EBIT to assets	Earnings before interest and taxes `ebit` divided by total assets `at`
ebitda_sale[b c]	EBITDA to sales	`ebitda` divided by total sales `sale`
gpm	Gross profit margin	Gross profit `gp` divided by total sales `sale`
gprof	Gross profit as % of assets	Gross profit `gp` divided by total assets `at`
ni_at	Net income to assets	Net income `ni` divided by tot. assets `at`
ni_mta[a d]	Net income to market assets	Net income `ni` divided by market assets `lt + prcc x csho`
npm	Net profit margin	Income before extraordinary items `ib` divided by total sales `sale`
opmad	Operating profit margin after depreciation	Operating income after depreciation `oiadp` divided by total sales `sale`
opmbd[b]	Operating profit margin before depreciation	Operating income before depreciation `oibdp` divided by total sales `sale`
pretret_noa	Pre-tax return on net operating assets	Operating income after depreciation `oiadp` divided by net operating assets, which equal net property, plant and equipment plus working capital `ppent+act-lct`
pretret_earnat	Pre-tax return on total earning assets	Operating income after depreciation `oiadp` divided by total earning assets, namely net property, plant and equipment plus current assets `ppent+act`
ptpm	Pre-tax profit margin	Pretax profit `pi` divided by total sales `sale`
re_at[b]	Retained earnings as % of assets	Retained earnings (`re`) divided by tot. assets (`at`)
roa	Return on assets	Operating income before depreciation `oibdp` divided by the average total assets `at` in the current and previous year.
roce	Return on capital employed	Earnings before interest and taxes `ebit` divided by the average capital employed in the current and previous year, whereby capital employed is given by the sum of total debt and common equity `dltt+dlc+ceq`.
roe	Return on equity	Income before extraordinary items `ib` divided by the average book equity `seq` in the current and previous year.
vola_prof[c]	Volatility of profitability	The standard deviation of the profitability ratio GProf in the years $t-4$, $t-3$, $t-2$, $t-1$, t
Activity ratios		
at_turn	Asset turnover	Total sales `sale` divided by the average tot. assets `at` in the current and previous year.
pay_turn	Payables turnover	Total sales `sale` divided by the average accounts payable `ap` in the current and previous year.
Valuation ratios		
bm	Book to market ratio	Book equity `seq` divided by market equity `prcc*csho`
div_payer[b]	Flag for dividend payer	Binary flag indicating if dividends `dvpsx_f` is positive
evm	Enterprise value multiple	The sum of total debt , market equity, preferred stock and balance sheet minority interest (`dltt+dlc+prcc*csho+pstk+mib`) divided by operating income before depreciation `oibdp`
mb[a b]	Market to book ratio	Market equity `prcc*csho` divided by book equity `seq`
ps	Price to sales ratio	Closing share price `prcc` divided by sales `sale`
pcf	Price to cash flow ratio	Closing share price `prcc` divided by operating cash-flow `ocf`
Market and size ratios		
BETA[b]	Systematic risk	Market model beta which is the coefficient of the CRSP value-weighted index return in the regression used to define SIGMA.
EXRET[a d]	Excess return	Annualized average excess log return over the NYSE, AMEX and ARCA index (using daily price data)
RSIZE	Relative size	Ratio of firm market capitalization to the total capitalization of NYSE, AMEX and ARCA index
lAT[c]	Size variable	Log of book assets (in 2005 dollars)
NYSE_SIZE_PERC[b]	Size variable	NYSE market capitalization percentile (the fraction of NYSE firms with capitalization smaller than or equal to firm i in year t)
PRICECAP[a d]	Capped stock price	Average daily stock price from CRSP on one calendar year basis capped at 15$
SIGMA[a c d]	Idiosyncratic risk	The root mean squared error from a regression of a firm's daily stock returns on the CRSP value-weighted index return. One firm-year observation of idiosyncratic risk is computed using firm-specific daily stock returns going back one year from the current date. A minimum of 52 observations in a year are required to calculate idiosyncratic risk.
Miscellaneous ratios		
accrual	Accruals to average assets	Operating activities net cash flow `oancf` minus income before extraordinary items `ib` divided by average tot. assets `at` in the current and previous year.

(*continued on next page*)

Table A.1 (*continued*).

Code	Ratio	Description
Capitalization ratios		
adv_sale	Advertising as % of sales	Advertising expenses `xad` to total sales `sale`
capx_at[b,c]	Capital expenditures to assets	Capital expenditures `capx` divided by total assets
ppent_at[b,c]	Tangibility	Net property, plant and equipment `ppent` divided by total assets
rd_at[b]	Research and development as % of assets	R&D expenses `xrd` divided by total assets
rd_sale	R&D as % of sales	R&D expenses `xrd` to total sales `sale`
sale_invcap	Sales per $ of invested capital	Total sales `sale` divided by `icapt`
sale_equity	Sales per $ of stockholders equity	Total sales `sale` divided by total equity `seq`
staff_sale	Labor expense as % of sales	Labor expenses `xlr` to total sales `sale`
xrent_at[c]	Rental expenses as % of assets	Rental expenses `xrent` divided by total assets

[a]Ratio in Campbell et al. (2008).
[b]Ratio in Alp (2013).
[c]Ratio in Baghai et al. (2014).
[d]Ratio in Tian et al. (2015).

Table A.2
This table displays the average value of the 13 ratios used in the analysis for each year of the sample period.

	BETA	che_mta	convdebt at	debt ebitda	div_payer	dltt_at	ebit_at	equity invcap	log(RSIZE)	lt_mta	NYSE SIZE PERC	re_at	SIGMA
1985	0.975	0.057	0.029	2.566	0.828	0.240	0.096	0.613	−8.099	0.497	0.626	0.263	0.018
1986	0.973	0.057	0.034	2.760	0.812	0.258	0.085	0.594	−8.218	0.474	0.618	0.255	0.020
1987	1.085	0.056	0.040	2.549	0.781	0.269	0.091	0.570	−8.161	0.441	0.642	0.242	0.024
1988	0.999	0.059	0.038	2.974	0.797	0.270	0.104	0.563	−8.093	0.494	0.655	0.243	0.019
1989	0.993	0.052	0.035	3.003	0.777	0.283	0.102	0.540	−8.081	0.485	0.664	0.226	0.018
1990	1.023	0.051	0.036	2.556	0.781	0.283	0.096	0.536	−8.084	0.557	0.682	0.221	0.022
1991	1.006	0.045	0.032	2.724	0.789	0.278	0.088	0.547	−8.095	0.498	0.694	0.226	0.021
1992	1.033	0.043	0.034	3.288	0.717	0.279	0.087	0.538	−8.233	0.496	0.685	0.194	0.021
1993	0.914	0.038	0.034	3.211	0.692	0.282	0.088	0.516	−8.261	0.470	0.694	0.155	0.020
1994	1.006	0.039	0.034	2.821	0.678	0.281	0.096	0.515	−8.242	0.472	0.715	0.145	0.019
1995	0.869	0.033	0.028	2.593	0.670	0.276	0.106	0.529	−8.332	0.456	0.713	0.159	0.019
1996	0.821	0.033	0.022	2.725	0.660	0.281	0.106	0.534	−8.455	0.434	0.709	0.163	0.020
1997	0.773	0.030	0.020	2.371	0.650	0.283	0.108	0.545	−8.561	0.391	0.708	0.173	0.020
1998	0.865	0.036	0.019	2.741	0.614	0.306	0.100	0.515	−8.774	0.474	0.694	0.169	0.027
1999	0.522	0.031	0.018	3.092	0.610	0.320	0.092	0.500	−9.017	0.479	0.684	0.164	0.029
2000	0.532	0.031	0.019	2.919	0.592	0.313	0.103	0.505	−9.199	0.501	0.673	0.165	0.034
2001	0.782	0.040	0.018	3.120	0.573	0.304	0.091	0.521	−8.915	0.507	0.682	0.175	0.028
2002	0.931	0.048	0.024	3.310	0.543	0.302	0.084	0.523	−8.732	0.506	0.688	0.160	0.027
2003	0.978	0.048	0.026	3.340	0.535	0.292	0.088	0.528	−8.732	0.469	0.677	0.148	0.020
2004	1.169	0.050	0.025	2.858	0.576	0.276	0.102	0.554	−8.606	0.424	0.696	0.162	0.017
2005	1.209	0.050	0.021	2.549	0.604	0.264	0.108	0.575	−8.563	0.397	0.699	0.174	0.017
2006	1.202	0.049	0.020	2.590	0.611	0.266	0.113	0.574	−8.550	0.406	0.701	0.178	0.017
2007	1.060	0.049	0.022	2.450	0.634	0.279	0.106	0.553	−8.530	0.405	0.700	0.194	0.018
2008	1.151	0.058	0.020	2.492	0.624	0.283	0.102	0.546	−8.565	0.487	0.687	0.190	0.034
2009	1.312	0.067	0.021	3.077	0.617	0.302	0.061	0.506	−8.585	0.486	0.689	0.145	0.029
2010	1.252	0.077	0.017	2.378	0.578	0.289	0.098	0.520	−8.511	0.462	0.698	0.144	0.017
2011	1.280	0.077	0.015	2.552	0.582	0.284	0.103	0.532	−8.513	0.499	0.699	0.184	0.018
2012	1.308	0.064	0.013	2.688	0.589	0.295	0.093	0.511	−8.636	0.473	0.686	0.167	0.018
2013	1.220	0.058	0.011	3.054	0.617	0.307	0.091	0.502	−8.531	0.424	0.704	0.162	0.015
2014	1.225	0.056	0.010	2.922	0.631	0.312	0.092	0.509	−8.584	0.419	0.701	0.167	0.017
2015	1.126	0.058	0.008	2.978	0.647	0.329	0.070	0.478	−8.640	0.464	0.703	0.155	0.019
2016	1.316	0.056	0.007	3.374	0.655	0.358	0.070	0.433	−8.699	0.442	0.693	0.117	0.021

Table A.3
This table displays the estimated regression coefficients for the model with equal coefficients (logit link).

Ratio	Estimate	SE	p value
BETA	−0.192	0.014	0.000
che_mta	−1.440	0.065	0.000
convdebt_at	−0.545	0.139	0.000
debt_ebitda	−0.005	0.000	0.000
div_payer	0.591	0.024	0.000
dltt_at	−1.955	0.093	0.000

(*continued on next page*)

Table A.3 (*continued*).

Ratio	Estimate	SE	p value
ebit_at	1.122	0.038	0.000
equity_invcap	0.534	0.066	0.000
log(RSIZE)	320.396	23.671	0.000
lt_mta	−0.265	0.064	0.000
NYSE_SIZE_PERC	4.297	0.091	0.000
re_at	0.307	0.040	0.000
SIGMA	−22.244	0.461	0.000

Table A.4
This table displays the estimated regression coefficients for the model with equal coefficients, where the previous ratings at time $t-1$ are also included as dummy variables in the model (logit link). The reference class for the dummy variables is the case where there was no rating observed in the previous year.

Ratio	Estimate	SE	p value
lagged rating dummy CCC	−5.593	0.196	0.000
lagged rating dummy B	−3.259	0.081	0.000
lagged rating dummy BB	−1.015	0.063	0.000
lagged rating dummy BBB	1.421	0.076	0.000
lagged rating dummy A	4.666	0.090	0.000
lagged rating dummy AA	8.776	0.155	0.000
lagged rating dummy AAA	13.712	0.230	0.000
BETA	−0.097	0.059	0.104
che_mta	−1.682	0.506	0.001
convdebt_at	−0.422	0.501	0.399
debt_ebitda	−0.005	0.004	0.211
div_payer	0.444	0.083	0.000
dltt_at	−2.453	0.451	0.000
ebit_at	3.252	0.364	0.000
equity_invcap	0.417	0.327	0.202
log(RSIZE)	96.453	50.455	0.056
lt_mta	−0.705	0.230	0.002
NYSE_SIZE_PERC	4.831	0.277	0.000
re_at	−0.524	0.140	0.000
SIGMA	−44.614	2.432	0.000

Table A.5
This table presents in-sample model performance metrics weighted Cohen's κ_w and the percentage of correct predictions within $k = 0, 1, 2, 3, 6$ rating notches for conditional and marginal predictions for both the proposed longitudinal model and for the longitudinal model which includes the lag-1 ratings as covariates (logit link). Note that the marginal predictions for the model with lagged ratings are better, given that this model includes information on the previous rating history on the covariate side.

Measure	LASSO	
	Longitudinal	Longitudinal (with lagged ratings)
Conditional predictions		
κ_w	0.9667	0.957
% of correct predictions	55.64%	53.24%
within $k = 1$ notches	92.80%	88.65%
within $k = 2$ notches	98.48%	97.50%
within $k = 3$ notches	99.61%	99.22%
within $k = 6$ notches	99.97%	99.93%
Marginal predictions		
κ_w	0.8231	0.9017
% of correct predictions	25.13%	29.12%
within $k = 1$ notches	57.61%	71.91%
within $k = 2$ notches	77.49%	91.52%
within $k = 3$ notches	91.21%	97.61%
within $k = 6$ notches	99.63%	99.92%

References

Afonso, António, Gomes, Pedro, Rother, Philipp, 2009. Ordered response models for sovereign debt ratings. Appl. Econ. Lett. 16 (8), 769–773. http://dx.doi.org/10.1080/13504850701221931.
Alp, Aysun, 2013. Structural shifts in credit rating standards. J. Finance 68 (6), 2435–2470. http://dx.doi.org/10.1111/jofi.12070.
Altman, Edward I., Kao, Duen Li, 1992. The implications of corporate bond ratings drift. Financ. Anal. J. 48 (3), 64–75. http://dx.doi.org/10.2469/faj.v48.n3.64.

Altman, Edward, Katz, Steven, 1976. Statistical bond rating classification using financial and accounting data. In: Proceedings of the Conference on Topical Research in Accounting. New York University Press New York, pp. 205–239.

Baghai, Ramin P., Servaes, Henri, Tamayo, Ane, 2014. Have rating agencies become more conservative? Implications for capital structure and debt pricing. J. Finance 69 (5), 1961–2005. http://dx.doi.org/10.1111/jofi.12153.

Bergmeir, Christoph, Hyndman, Rob J., Koo, Bonsoo, 2018. A note on the validity of cross-validation for evaluating autoregressive time series prediction. Comput. Statist. Data Anal. 120, 70–83. http://dx.doi.org/10.1016/j.csda.2017.11.003.

Blume, Marshall E., Lim, Felix, Mackinlay, A. Craig, 1998. The declining credit quality of U.S. corporate debt: Myth or reality? J. Finance 53 (4), 1389–1413. http://dx.doi.org/10.1111/0022-1082.00057.

Campbell, John Y., Hilscher, Jens, Szilagyi, Jan, 2008. In search of distress risk. J. Finance 63 (6), 2899–2939. http://dx.doi.org/10.3386/w12362.

Cohen, Jacob, 1968. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. Psychol. Bull. 70 (4), 213–220. http://dx.doi.org/10.1037/h0026256.

Creal, Drew, Schwaab, Bernd, Koopman, Siem Jan, Lucas, Andre, 2014. Observation-driven mixed-measurement dynamic factor models with an application to credit risk. Rev. Econ. Stat. 96 (5), 898–915. http://dx.doi.org/10.1162/REST_a_00393.

Dimitrakopoulos, Stefanos, Kolossiatis, Michalis, 2016. State dependence and stickiness of sovereign credit ratings: evidence from a panel of countries. J. Appl. Econometrics 31 (6), 1065–1082. http://dx.doi.org/10.1002/jae.2479.

Ederington, Louis H., 1985. Classification models and bond ratings. Final. Rev. 20 (4), 237–262. http://dx.doi.org/10.1111/j.1540-6288.1985.tb00306.x.

Friedman, Jerome, Hastie, Trevor, Tibshirani, Rob, 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33 (1), 1. http://dx.doi.org/10.18637/jss.v033.i01.

Gentry, James A., Whitford, David T., Newbold, Paul, 1988. Predicting industrial bond ratings with a probit model and funds flow components. Final. Rev. 23 (3), 269–286. http://dx.doi.org/10.1111/j.1540-6288.1988.tb01267.x.

Hastie, Trevor, Tibshirani, Robert, Friedman, Jerome, 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Science & Business Media.

Hilscher, Jens, Wilson, Mungo, 2016. Credit ratings and credit risk: Is one measure enough? Manage. Sci. 63 (10), 3414–3437. http://dx.doi.org/10.1287/mnsc.2016.2514.

Hirk, Rainer, Hornik, Kurt, Vana, Laura, 2019. Multivariate ordinal regression models: An analysis of corporate credit ratings. Stat. Methods Appl. 507–539. http://dx.doi.org/10.1007/s10260-018-00437-7.

Hirk, Rainer, Hornik, Kurt, Vana, Laura, 2020. **mvord**: An r package for fitting multivariate ordinal regression models. J. Stat. Softw. 93 (4), 1–41. http://dx.doi.org/10.18637/jss.v093.i04.

Horrigan, James O., 1966. The determination of long-term credit standing with financial ratios. J. Account. Res. 44–62. http://dx.doi.org/10.2307/2490168.

Hwang, Ruey-Ching, 2013. Predicting issuer credit ratings using generalized estimating equations. Quant. Finance 13 (3), 383–398. http://dx.doi.org/10.1080/14697688.2011.593542.

Hwang, Ruey-Ching, Chung, Huimin, Chu, C.K., 2010. Predicting issuer credit ratings using a semiparametric method. J. Empir. Financ. 17 (1), 120–137. http://dx.doi.org/10.1016/j.jempfin.2009.07.007.

Kaplan, Robert S., Urwitz, Gabriel, 1979. Statistical models of bond ratings: A methodological inquiry. J. Bus. 231–261, URL https://www.jstor.org/stable/2352195.

Manso, Gustavo, 2013. Feedback effects of credit ratings. J. Financ. Econ. (ISSN: 0304-405X) 109 (2), 535–548. http://dx.doi.org/10.1016/j.jfineco.2013.03.007.

McFadden, Daniel, 1977. Quantitative Methods for Analyzing Travel Behavior of Individuals: Some Recent Developments. Institute of Transportation Studies, University of California.

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

Paulo, Alves, 2018. Abnormal retained earnings around the world. J. Multinatl. Final. Manag. 46, 63–74. http://dx.doi.org/10.1016/j.mulfin.2018.05.002.

Pinches, George E., Mingo, Kent A., 1973. A multivariate analysis of industrial bond ratings. J. Finance 28 (1), 1–18. http://dx.doi.org/10.2307/2978164.

Pogue, Thomas F., Soldofsky, Robert M., 1969. What's in a bond rating. J. Financ. Quant. Anal. 201–228. http://dx.doi.org/10.2307/2329840.

Poon, Winnie P.H., Firth, Michael, Fung, Hung-Gay, 1999. A multivariate analysis of the determinants of Moody's bank financial strength ratings. J. Int. Final. Mark. Inst. Money 9 (3), 267–283. http://dx.doi.org/10.1016/S1042-4431(99)00011-6.

R Core Team, 2022. R: A Language and Environment for Statistical Computing. URL https://www.R-project.org/.

Reusens, Peter, Croux, Christophe, 2017. Sovereign credit rating determinants: A comparison before and after the European debt crisis. J. Bank. Financ. 77, 108–121. http://dx.doi.org/10.1016/j.jbankfin.2017.01.006.

Sermpinis, Georgios, Tsoukas, Serafeim, Zhang, Ping, 2018. Modelling market implied ratings using LASSO variable selection techniques. J. Empir. Financ. 48, 19–35. http://dx.doi.org/10.1016/j.jempfin.2018.05.001.

Tian, Shaonan, Yu, Yan, Guo, Hui, 2015. Variable selection and corporate bankruptcy forecasts. J. Bank. Financ. 52, 89–100. http://dx.doi.org/10.1016/j.jbankfin.2014.12.003.

Tibshirani, Robert, 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B Stat. Methodol. 58 (1), 267–288. http://dx.doi.org/10.1111/j.1467-9868.2011.00771.x.

Tuzcuoğlu, Kerem, 2019. Composite Likelihood Estimation of an Autoregressive Panel Probit Model with Random Effects. Bank of Canada Staff Working Paper, http://dx.doi.org/10.2139/ssrn.3381994, 16.

Varin, Cristiano, 2008. On composite marginal likelihoods. AStA Adv. Stat. Anal. 92 (1), 1. http://dx.doi.org/10.1007/s10182-008-0060-7.

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

West, Richard R., 1970. An alternative approach to predicting corporate bond ratings. J. Account. Res. 118–125. http://dx.doi.org/10.2307/2674717.

Wurm, Michael J., Rathouz, Paul J., Hanlon, Bret M., 2017. Regularized ordinal regression and the **ordinalnet** r package. Stat. URL http://arxiv.org/abs/1706.05003.