
There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

http://eprints.gla.ac.uk/164498/

Deposited on 25 June 2018

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk
Modelling market implied ratings using LASSO variable selection techniques

by

George Sermpinis, Serafeim Tsoukas, Ping Zhang
University of Glasgow

January 2018

Abstract
Making accurate predictions of corporate credit ratings is a crucial issue to both investors and rating agencies. In this paper, we investigate the determinants of market implied credit ratings in relation to financial factors, market-driven indicators and macroeconomic predictors. Applying a variable selection technique, the least absolute shrinkage and selection operator (LASSO), we document substantial predictive ability. In addition, when we compare our LASSO-selected models with the benchmark ordered probit model, we find that the former models have superior predictive power and outperform the latter model in all out-of-sample predictions.

Key words: Market implied ratings, LASSO, Financial ratios, Forecasting

JEL: G24, G33, C25, O16

\(^1\) We are grateful for comments and inputs from an anonymous referee, Oldrich Vasicek and from participants at the 2017 European Financial Management Association Conference, the 2017 European Meetings of the Econometric Society, the 2016 International Conference on Forecasting Financial Markets and the Adam Smith Business School seminars. Corresponding Author: Serafeim Tsoukas, email: serafeim.tsoukas@glasgow.ac.uk. Adam Smith Business School, University of Glasgow, G12 8QQ, Tel: +44 141 3306325. The third author is grateful to China Scholarship Council for funding (Grant number: 201508060273). Any remaining errors are our own.
1. Introduction

Recent turmoil in the financial markets has focused attention on the rating agencies and the process by which they assign ratings to firms and their financial obligations. A careful management of credit risk is high the agenda for both market participants and regulators. It is well known that the credit ratings of the top three nationally recognized statistical rating organizations (NRSROs), Standard and Poor’s, Moody’s and Fitch play a key role in the pricing of credit risk and in the delineation of investment strategies since they measure the firm’s long-term ability and willingness to meet debt servicing obligations. As such, the ratings indicate the probability that a given borrower will default. However, the accuracy and the timing of the ratings have been heavily criticized, especially during the most recent financial crisis. It has been argued that the standard agency ratings do not adjust quickly to price changes and therefore may be out of date. In response to these concerns, Fitch has recently developed a new model to derive Market Implied Ratings (MIRs) from bond and equity prices. The obvious advantage of these ratings compared to the conventional agency ratings is that they adjust instantaneously to price changes.

This study offers methodological extensions applying a variable selection approach, the least absolute shrinkage and selection operator (LASSO), and its most promising derivation, the Elastic net, into ordered probit and continuation ratio models, to the task of predicting Fitch’s CDS and Equity implied ratings (CDSIRs and EQIRs respectively hereafter). The research aims to exploit the LASSO properties and unveil the underlying structure of CDSIRs and EQIRs. There are several studies that use accounting ratios and other publicly available information in reduced-form models in order to predict credit ratings. These studies use various techniques (OLS, multinomial and ordered logit/probit models) to identify the most important characteristics for predicting bond ratings (see for instance the early studies by Poldue and Soldowsky, 1969, Pinches and Mingo, 1973, Kaplan and Urwitz, 1979 and Kao and Wu, 1990). The upshot is that financial healthiness is associated with ratings determination and prediction of default. Another line of research advocates the importance of estimating the

---

2 With respect to the latter aim of this study, we do not report estimated coefficients of the prediction models to preserve space. These results are available upon request.
models in a dynamic setting and documents a noticeable improvement in the predictive ability of the models once state dependence is controlled for (see Mizen and Tsoukas, 2012). One drawback of the reduced-form models, discussed above, is that they tend to employ many rating predictors as inputs despite the fact that only a sub-set is relevant. This has two critical implications. First, this approach can omit potentially important determinants leading to a decrease in prediction accuracy. Second, given the large number of predictors included, it does not provide a sparse representation, implying that these models cannot be readily used by market participants and rating agencies.

Our approach is mostly related to the literature that examines the determinants of credit ratings, but we add to it in two important ways. First, we make a methodological contribution by deriving a simple and more intuitive, yet innovative model, which is based on the variable selection technique, pioneered by Tibshirani (1996)—the least absolute shrinkage and selection operator (LASSO). It is well accepted that this selection approach not only helps in identifying the most relevant predictors from an extensive set of candidate variables, but also improves the predictive power (see Fan and Li, 2001 and Tian et al, 2015). In addition, LASSO does not require strict assumptions such as a pre-selection of the variables considered and it is consistent statistically as the number of observations approach infinity (Van de Geer, 2008). Importantly, LASSO can potentially sidestep the problem of multicollinearity, which is fairly common in probit/logit models, and it is computationally efficient even when considering a large set of potential predictors. Our study is the first, as far as we know, to provide a systematic empirical analysis of LASSO in MIRs forecasts. In doing so, we explore the relative importance of several time-varying covariates from an extensive set of firm-, market-specific and macro-economic explicators used to predict market implied ratings. This is important as we provide a parsimonious set of predictors that can be readily implemented by investors, managers and credit risk agencies.

Second, we use a data-set made up by MIRs instead of the standard long-term ratings. The former type of ratings represents an innovation to the ratings industry to address the issue of staleness in their long-term counterparts. The market implied ratings rely on proprietary and data-intensive rating models that incorporate market information into a model-based credit
assessment (see for instance, Rösch, 2005 and Tsoukas and Spaliara, 2014). The most appealing characteristic of these ratings is that they can adjust instantly to market changes. Hence, we build on the foundations of the literature on implied (or point-in-time ratings) by investigating the forecasting power of models that capture volatile market changes.

To preview our findings, we show that several financial factors along with market-driven and macroeconomic variables contain information about market implied ratings. Importantly, when applying the LASSO techniques, we are able to significantly improve the predictive power of our models in out-of-sample predictions compared to the ordered probit model, which is commonly adopted in the literature. Moreover, we note that the LASSO models with BIC-type tuning parameter selector outperform their LASSO counterparts with AIC-type selector for the dataset and periods under study. Thus, LASSO-selected models display improved forecasting power.

The rest of the paper is laid out as follows. In section 2 we discuss the relevant literature. Section 3 presents the data and summary statistics. In section 4 we describe our methodology. In Section 5 we report the empirical results and robustness tests. Section 6 concludes the paper.

2. Related literature

The issue of how rating agencies use public information in setting quality ratings has attracted considerable attention in the literature. In fact, the literature goes as far back as Horrigan (1966). This study presents the first attempt to predict ratings based on the characteristics of the bonds and the issuing firms. The author concentrates on accounting data and financial ratios in order to find the most appropriate predictors. The set of preferred variables contains total assets, net worth to total debt, net operating profit to sales and working capital. Poldue and Soldofsky (1969) also assign ordinal numbers to ratings and investigate different accounting variables as potential determinants. They conclude that the most significant independent variables are long-term debt to total assets, the coefficient of variation of
earnings, and total assets. West (1970) challenge Horrigan’s study by using another set of explanatory variables namely earnings volatility, capital structure, reliability and marketability. Based on values of the obtained R-squared, the author claims that the proposed model has a better explanatory power.

Pinches and Mingo (1973) adopt a two-stage approach to assign ratings to bond issues. This study attempts to test the predicting ability of a small number of explanatory variables using multiple regression and discriminant analysis. The proportion of correct predictions lie in the region of 70 percent. Moreover, Kaplan and Urwitz (1979) confirm the above studies by using an ordered probit analysis. They show that ratings may be reasonably well predicted using balance sheet information. Other studies that use a small number of explanatory variables (e.g., leverage, profitability, interest coverage, firm’s size and subordination status) to predict credit ratings include Ederigton (1985) and Gentry et al. (1988). The former study uses long-term debt, subordination, total assets and interest coverage as explanatory variables, while the later focuses on subordination, size of issue, debt ratio, cumulative years that dividends were paid and net income to interest.

Blume et al. (1998) depart from the traditional examination of credit ratings determinants by considering whether there is any tendency for a company that maintains the same values of accounting ratios over time to receive a lower rating due to worsening of rating standards. Using an ordered probit analysis they find that rating agencies have changed the way in which they evaluate credit standing and they report a secular tightening of rating agency standards. They conclude that it became more difficult for firms to obtain improved ratings in the mid-1980s and early 1990s.

More recently, ordered probit methodologies are employed by Hwang et al., (2009) and Hwang et al., (2010) to forecast credit ratings. Both studies show that several predictors are important in forecasting credit ratings such as the size of the company, balance sheet position, stock market performance and industry effects. In addition, modelling long-term ratings in a dynamic setting has shown improvements in forecasting (see Hwang, 2011 and Hwang, 2013).

---

3 The study shows that both profitability and coverage ratio are insignificant and quantitatively unimportant.
In a similar vein, Mizen and Tsoukas (2012) find that allowing for persistence in ratings significantly improves the forecasting power of long-term ratings. In a subsequent study, Tsoukas and Spaliara (2014) use the market implied ratings and the ordered probit modelling strategy to investigate the role of financial constraints. They conclude that financial variables are more important in predicting credit ratings for firms likely to face financing constraints.

The literature on market implied ratings has focused on the comparison between long-term agency ratings and market implied ratings. Breger et al. (2003) use bond spreads to find suitable thresholds categorizing bonds. They find that implied ratings are a superior application to identify default probability in the rating system. Rösch (2005) documents that implied ratings can provide more accurate default probability forecasts than long-term ratings. Castellano and Giacometti (2012) note that MIRs can be regarded as early warning signals of credit rating changes.

Moving to the line of work on variable selection techniques, there is only a handful of papers investigating default probabilities in various settings. Härdle and Prastyo (2013) employ the LASSO approach to predict default probability in a sub-set of Asian economies. Amendola et al., (2012) evaluate the default risk in the limited liability sector in Italy. Finally, Tian et al., (2015) evaluate the probability of bankruptcy using a comprehensive sample of US firms. The authors conclude that the accuracy in the out-of-sample prediction can be superior to previous studies of estimating default by combining reduced-form models with the LASSO procedure.

The studies discussed above provide a useful background on the credit ratings procedure and the selection process of relevant predictors. In the sections that follow we turn to our data and estimation strategy.

3. Data and summary statistics

3.1 Data sources

The data on market implied ratings are taken from Fitch’s database and refer to solicited ratings for all traded US corporations. This database provides information on the CDS and
Equity implied ratings assigned to each issuer as well as the date that the rating became available. Both CDS and Equity implied ratings are reported on a monthly frequency and span the period 2002 to 2008. In keeping with the normal practice in the literature, we categorize our firms into rating buckets without consideration of notches (i.e. + or -). Amato and Furfine (2004) and Mizen and Tsoukas (2012) note that this classification takes into account large cumulative changes of ratings rather than small movements notch by notch, and avoids generation of rating categories with very few observations. Therefore, we consider seven rating categories, ranging from AAA to Below CCC, which are assigned numerical values in Table 1, starting with 1 to AAA, 2 to AA..., 7 to Below CCC.

Firm-specific accounting data are extracted from Fitch’s Peer Analysis Tool. Corporate historical data for all firms rated by Fitch are available on a quarterly basis from this database. For these firms with credit ratings, we link their ratings to Fitch’s balance sheet statements and profit and loss accounts. Hence, our dataset is constructed by merging monthly market implied ratings data and quarterly firm-level accounting data. In other words, we have an entry for each firm-month with CDSIRs and EQIRs data and financial and market data. Following commonly used selection criteria in the literature; we exclude companies that do not have complete records on our explanatory variables and firm-months with negative sales and assets. To control for the potential influence of outliers, we winsorize the regression variables at the 1st and 99th percentiles.

Data on market indicators and macroeconomic variables are sourced from Bloomberg. These data items are reported monthly. Our combined sample contains data for 211 firms that

---

4 The research aims to study the structure and the predictability of the implied ratings in the years preceding the recent global financial crisis. Our choice to focus on a time window ending in 2008 is motivated by two important considerations. First, the global financial crisis and the collapse of Lehman Brothers constitute a shock that may have acted as a confounding factor to the determination of credit ratings. In fact, it can be argued that the misinterpretation of the credit ratings by investors was one of the main contributors to the crisis. Second, the data were downloaded early in 2010 owing to a research project supported by Fitch: the coverage period is therefore 2002–2008.

5 In EQIRs we do not observe any ratings in the last category, hence six groups are generated for this type of implied ratings.

6 For each firm, the quarterly data is repeated every month during the same quarter if the monthly data is not observed.
operate in all sectors of the US economy except agriculture, forestry and fishing and public administration. The panel has an unbalanced structure with the number of observations on each firm varying between 1 and 63. Our sample presents two characteristics that make it especially appealing for our analysis. First, it includes both investment and speculative grades ratings in line with previous studies (see for instance Amato and Furfine, 2004). This is particularly beneficial since we can cover the entire spectrum of firms. Second, the distribution of agency (long-term) ratings in CDS data is very similar to the distribution of agency ratings in the general bond population (see Fitch 2007). Thus, both the CDSIRs and the EQIRs databases can provide a representative base for conducting our empirical analysis.

3.2 Choice of explanatory variables

Prior empirical research on the determinants of credit ratings has considered both business and financial risks. The former type of risk includes an assessment of industry characteristics, firm size, management capability and organizational factors. The latter concerns the quality of a firm’s accounting procedures, profitability, cash flow situation and its overall financial policy. In the market implied ratings, models typically consider market-related information in addition to the above-mentioned factors. With that in our mind, we also turn to rating agencies and in particular to Fitch, to find out what matters when assigning a market implied rating. In other words, the selection of our explanatory variables is guided both from the existing empirical literature (see for example Kaplan and Urwitz, 1979; Ederington, 1985; Poon, 2003; Chava and Jarrow, 2004; Amendola et al, 2011; Mizen and Tsoukas, 2012; Hwang, 2013; Creal et al, 2014; Doumpos et al, 2015 and Tian et al, 2015), and the common practice of rating agencies (see Fitch 2007 and Liu et al. 2007).7

3.2.1 Firm-specific variables

We consider 16 firm-specific accounting variables as potential predictors of ratings. These variables are intended to measure different aspects of firms’ financial health, these are size, leverage, coverage, cash flow, profitability and liquidity. Specifically, we employ the firm size

7 The expected relationship between these variables and MIRs is presented in Table A1 in on-line Appendix A. The Table also provides a detailed description of the variables used in this study.
(DETA) as measured by the natural logarithm of firms’ real total assets. Size accounts for the scale of the firm and would be expected to improve the rating. Next, we proxy leverage using a number of ratios: The ratio of long-term debt to total assets (LDA), the ratio of short-term debt to total assets (SDA), the ratio of total debt over total assets (TDA), the ratio of total assets over equity (AE), and the ratio of total debt to earnings before interest, taxes, depreciation, amortization, and restructuring or rent costs (TDEBITDA). Higher values of the above ratios are likely to increase financial risk and hence should worsen the rating. The next two measures capture the creditworthiness of the firm as they show the firm’s ability to generate income to meet interest rate obligations: the ratio of earnings before interest and tax over interest expenses (EBITINT) and the ratio of total debt to earnings before interest, taxes, depreciation, amortization, and restructuring or rent costs to interest expenses (EBITDAINT). Both ratios would improve the credit rating if they were to increase. Further, cash flow is measured by the following ratios: cash flow from operating activities over total assets (CFOA), and cash and equivalent over total assets (CASHEQA). We expect firms with higher cash flows to have improved ratings. The following five ratios measure firm profitability: The ratio of operating income to net sales (OM), the ratio of net income without dividends over total capital (ROC), the ratio of net income over shareholders’ equity (ROE), the ratio of net income over total assets (ROA) and the ratio of the funds from operations to total debt (FFD). An increase in the above-mentioned profitability ratios should be associated with an improvement of ratings. Finally, liquidity is measured by the ratio of cash from operations to liabilities (LIQ), which indicates a firm’s ability to satisfy its short-term obligations as they become due. Higher levels of liquidity should improve credit ratings.

3.2.2 Market-driven indicators

As noted above, market implied ratings are likely to be determined by market-related conditions. Therefore, we employ the following market indicators: Excess return (EXRET) as measured by the monthly stock return on the firm minus the S&P 500 index return. The relative size of a firm in market (FSIZE) measured by each firm’s market equity value divided by the total market equity value. The above-mentioned variables should be positively correlated with ratings upgrades. Next, we use the volatility of stock return (STD) which is calculated as the standard deviation of each company’s monthly stock returns. We also use
the systematic risk of each firm (BETA), measured by the Capital Asset Pricing Model for each firm. Finally, we extract the 1-year and 5-year default probabilities (PD1 and PD5) from the Fitch’s Peer Analysis Tool. All three variables should worsen ratings if they were to increase.

3.2.3 Macroeconomic influences

We also consider an extensive list of macro-economic covariates as potential predictors of market implied ratings. Specifically, the stock market performance is evaluated by the S&P 500 return, which calculates returns on the S&P 500 index (RLSP). The short-term interest rate as measured by the three-month commercial paper rate (CPFFM), three-month Treasury bill rate minus federal funds rate (TB3) and the one-year constant maturity treasury rate (GS1). We also employ the general price level, as measured by the growth rate in the narrow money stock (MB) and inflation rate (INFL). The aggregate economic activity is captured by the rate of change in industrial production (DLIP), the index of the growth rate of real GDP (DLDGD), the average of monthly Chicago Fed National Activity Index over the year (CFNA), the average of monthly unemployment rate over the year (UNRATE) and the Chicago Board Options Exchange (CBOE) volatility index (VIX). All macro-variables, apart from VIX, are reported in percentages. The eleven macroeconomic variables measure different aspects of the aggregate economy’s performance. Their relationship with the market implied ratings could be either positive or negative as ratings tend to improve during good times but agencies have been observed to tighten their standards during these periods. Hence, the relationship between ratings and macro-economic variables is an issue that will be determined empirically.

3.3 Summary statistics

Tables 2 and 3 illustrate the distribution of firms by rating category for CDSIRs and EQIRs, respectively. It can be observed that the distribution of firms across the rating categories is quite stable and that most companies are assigned A and BBB ratings.

Insert Table 2

Insert Table 3
At the next stage, we report summary statistics for our explanatory variables in Tables 4 and 5. We present statistics splitting the sample between investment grade and non-investment grade to gauge any differences across ratings categories. P-values for the tests of equality of means across the above-mentioned groups are reported in the last columns of the tables. We observe, as expected, that firms in the investment grade group display better financial characteristics, as measured by the balance sheet indicators. The tests point to significant differences between the two groups, which indicate that there is a correlation between better financial health and an improved rating. In other words, there is considerable cross-sectional variation in market implied ratings. Moving to the market indicators, we find that improved market conditions are associated with investment grade ratings, which also suggests a link between the market climate and the ratings.

Insert Table 4
Insert Table 5

4. Methodology

We predict the changes in market implied ratings with ordered probit (OP) and continuation ratio (CR) models combined with LASSO or the Elastic net. The proposed methodology aims at selecting the most important predictors and at providing accurate MIRs forecasts. LASSO, originally proposed by Tibshirani (1996), is an extended form of an OLS regression which performs both variable selection and regularization through a shrinkage factor. It is capable of enhancing the accuracy and the interpretability of classical regression methods (Tibshirani, 1996). To maintain the properties of LASSO and capture the ordinal ranking of MIRs, penalty functions from LASSO or its variant (Elastic net) are added into OP or CR models. This helps us unveil the relation between the potential predictors (at the firm and macro level) and identify their significance in predicting MIRs. As benchmark to our study, we rely to the standard OP model. A description of the empirical modelling strategy follows.

4.1 OP

MIRs as a branch of credit ratings are discrete-valued signs and have an ordinal ranking. To meet the ordinal property of MIRs, OP is applied naturally as a benchmark in the relevant literature (Kaplan and Urwitz, 1979; Gentry, 1988; Blume et al., 1998; Amato and Furfine,
OP takes into account both the existence of ordinal ranking and the difference between any two adjacent ratings. We define the categorical variable $y_{it} = 1, 2, ..., 7$ according to the rating assigned to each firm. We assume that there is an unobservable dependent variable $y_{it}^*$ associated with $y_{it}$, which can be expressed as:

$$y_{it}^* = X_{it}\beta_1 + X_{it-1}\beta_2 + X_{it-2}\beta_3 + X_{it-3}\beta_4 + W_{it}\beta_5 + W_{it-1}\beta_6 + W_{it-2}\beta_7 + W_{it-3}\beta_8 + Z_{it}\beta_9 + Z_{it-1}\beta_{10} + Z_{it-2}\beta_{11} + Z_{it-3}\beta_{12} + y_{it-1}\beta_{13} + \epsilon_{it},$$

where $i = 1, 2, ..., N$ represents firms, and $t = 1, 2, ..., T$ represents different time periods. In this context, $t$ is the month end for monthly data. $\beta$ are vectors of unknown parameters to be estimated. $X$ denotes a set containing 16 accounting variables, which can be divided into broad groups of size, leverage, coverage, cash flow, profitability and liquidity. $W$ and $Z$ contain 6 market-driven variables and 11 macroeconomic variables, respectively. Following the literature (Güttler and Wahrenburg, 2007), all predictors in $X$, $W$ and $Z$ are lagged three periods denoted by $t - 1$, $t - 2$ and $t - 3$ to mitigate potential time tendency. To capture potential non-linear influences, we allow for non-linear transformation of the variables and therefore the square of each predictor is considered and included. Thus, the total number of firm-specific accounting, market driven and macroeconomic variables is 264. $y_{it-1}$ is an indicator of the firm’s rating in the previous time periods. We consider 4 lags to control for state dependence. The concern of persistency in ratings is an important dimension of time-series variation and the use of models with lagged rating categories is the standard way of addressing this issue. The error term $\epsilon_{it}$ in equation (1) is assumed to be a normally distributed random variable.

---

8 For details on the exposition of the OP, see Maddala (2008, pp 47-48).
9 This approach is justified theoretically since some variables may have a positive effect up to a certain (turning) point and negative thereafter.
10 We consider 16 accounting variables, 6 market-driven variables, 11 macroeconomic variables and their first three lags (in total, 132 variables). We also consider the square of these variables, thus in equation (1) we have in total 268 variables made up by 264 firm-specific accounting, market driven and macroeconomic variables and 4 state dependence variables. We use the same set of predictors throughout the paper.
11 State dependence captures previous rating state and indicates the realization of a rating in the previous time period. Following Contoyannis et al. (2004) and Mizen and Tsoukas (2012), the state dependence is controlled for by applying dummy variables representing first lags of each category in the dependent variable. Given that we observe limited observations in the extreme high and low rating categories, these ratings in the state dependence variables are merged into 5 groups, such as above AA ratings, A rating, BBB rating, BB rating and below B ratings. In addition, to avoid the dummy variable trap we omit one baseline rating category. Therefore, the lagged rating category BBB is not included in the models because it is regarded as the baseline category of lagged MIRs.
distributed residual with a zero mean and unit variance. In our data $y_{it}^*$ is not observed. Thus what is observed are the market implied ratings assigned to firms, which can take $M$ values. The relation between the observed variable $y_{it}$ and the latent variable $y_{it}^*$ it is assumed to be given by:

$$y_{it} = m \text{ if } \alpha_{m-1} < y_{it}^* \leq \alpha_m \text{ for } m = 1, \ldots, M$$  \hspace{1cm} (2)

For a set of parameters $\alpha_0$ to $\alpha_M$, where $\alpha_0 < \alpha_1 < \ldots < \alpha_M$, $\alpha_0 = -\infty$ and $\alpha_M = \infty$. Assuming a standard normal distribution for $\epsilon_{it}$, the conditional probabilities can be derived as:

$$P(r(y_{it} = m) = \Phi(\alpha_m - X_{it}\beta_1 - X_{it-1}\beta_2 - X_{it-2}\beta_3 - X_{it-3}\beta_4 - W_{it}\beta_5 - W_{it-1}\beta_6 - W_{it-2}\beta_7 - W_{it-3}\beta_8 - W_{it-4}\beta_9 - W_{it-5}\beta_{10} - W_{it-6}\beta_{11} - W_{it-7}\beta_{12} - y_{it-1}\beta_{13} - \Phi(\alpha_{m-1} - X_{it}\beta_1 - X_{it-1}\beta_2 - X_{it-2}\beta_3 - X_{it-3}\beta_4 - W_{it}\beta_5 - W_{it-1}\beta_6 - W_{it-2}\beta_7 - W_{it-3}\beta_8 - W_{it-4}\beta_9 - W_{it-5}\beta_{10} - W_{it-6}\beta_{11} - W_{it-7}\beta_{12} - y_{it-1}\beta_{13}))$$  \hspace{1cm} (3)

Where $\Phi(.)$ is the standard normal distribution function. We can evaluate the above probabilities for any combination of parameters in the vectors $\alpha$ and $\beta$.

4.2 OP with LASSO

According to Tibshirani (1996), LASSO is a method of regression that enables estimation and variable selection simultaneously in a non-orthogonal setting. Under a suitable choice of penalty power, LASSO selects variables by forcing some coefficients to zero and shrinking others. This reduces the variance of the estimated value and increases the accuracy of the regression prediction. The LASSO estimator resolves the $l_1$-penalized OP problem of estimating $\beta$ by maximizing a likelihood function. In particular, the maximization of the likelihood proceeds subject to the constraint $\sum_{q=1}^{p}|\beta_q| \leq s$, where $s$ is a user-specified tuning parameter and $q = 1, 2 \ldots p$ indicates the number of surviving predictors with non-zero estimated coefficients. This penalty corresponds to the L1 norm, and therefore it is often referred to as L1 penalized model. The OP with LASSO (L1 penalized OP model) can then be expressed as:
\[ \hat{\beta} = \arg\max_{\beta} \left( \ell(\beta | y_{it}, A) - \lambda \sum_{q=1}^{p} |\beta_q| \right). \tag{4} \]

where \( \ell(\beta | y_{it}, A) \) is the likelihood function of OP and \( A \) contains the pool of the potential predictors \( (X_{it}, X_{it-1}, X_{it-2}, X_{it-3}, W_{it}, W_{it-1}, W_{it-2}, W_{it-3}, Z_{it}, Z_{it-1}, Z_{it-2}, Z_{it-3}, y_{it-1}) \). All explanatory variables are standardized before applying the LASSO estimator. In equation (4), \( \lambda \) stands for the tuning parameter. As \( \lambda \) increases the sum of absolute values of estimated coefficients is reduced and shrinkage of coefficients is achieved. If \( \lambda \) exceeds a threshold value in the corresponding models, some estimated coefficients are set to zero ultimately. This “L1 norm penalty” or the constraint formulation in LASSO generates a more interpretable and sparse model.

As already noted, compared with other independent variable selection methods, LASSO can provide more stable and restricted models (Tibshirani, 1996; Fan and Li, 2001; Zou, 2006 and Tian et.al., 2015). It is also a computationally simple and efficient method (Efron et al., 2004). Several approaches, such as cross validation and information criteria, have been proposed in selecting latent models with minimum prediction errors or maximum log-likelihood estimation. Zou et.al. (2007) provide an efficient approach to obtain the optimal LASSO fit with the Akaike information criterion (AIC) (Akaike, 1974) and the Bayesian information criterion (BIC) (Schwartz, 1978). Sun and Zhang (2012) note that the computational cost of applying cross-validation in penalized models is considerable, while the theory of applying cross-validation is poorly understood. Therefore, AIC and BIC are used in the present study to select the tuning parameter and further detect the “best” model among a series of candidate models in OP with LASSO\(^{12}\). The “best” model in variable selection procedure will be chosen as the one that achieves the minimum AIC or BIC value. The exact algorithm behind this process is presented in on-line Appendix C. These models are benchmarked with their LASSO 10-cross validation counterparts.

\(^{12}\) It is well-known that AIC and BIC have different properties in model selection (for details see Yang, 2005; Shao, 1997 and Zhang et al, 2010).
4.3 OP with Elastic net

Elastic net is a LASSO variant introduced by Zou and Hastie (2004) that can further improve the accuracy of the estimation in the presence of highly correlated predictors. The Elastic net allows “grouping” variables in the model by adding a $l_2$-penalty from ridge regression. The OP with Elastic net resolves the $l_1$-penalized and $l_2$-penalized OP problem of estimating $\beta$ to maximize the OP likelihood function. The maximization of the likelihood proceeds subject to the following constraints (penalty functions) $\sum_{q=1}^{p} |\beta_q| \leq s_1$ and $\sum_{q=1}^{p} (\beta_q^2) \leq s_2$, where $s_1$ and $s_2$ are user-specified tuning parameters. These penalty functions correspond to the L1 and L2 norm. The OP with Elastic net model is presented below:

$$
\hat{\beta} = \arg\max_{\beta} \left( \ell(\beta|y_{it}, A) - \lambda_1 \sum_{q=1}^{p} |\beta_q| - \lambda_2 \sum_{q=1}^{p} (\beta_q^2) \right).
$$

(5)

where $\ell(\beta|y_{it}, A)$ is the likelihood function of the OP and $A$ contains the pool of the potential predictors $(X_{it}, X_{it-1}, X_{it-2}, X_{it-3}, W_{it}, W_{it-1}, W_{it-2}, W_{it-3}, Z_{it}, Z_{it-1}, Z_{it-2}, Z_{it-3}, y_{it-1})$.

Equation (5) can be converted as follows:

$$
\hat{\beta} = \arg\max_{\beta} \left( \ell(\beta|y_{it}, A) - \lambda \sum_{q=1}^{p} \left( \alpha |\beta_q| + \frac{1}{2} (1 - \alpha) (\beta_q^2) \right) \right),
$$

(6)

where $0 \leq \alpha \leq 1$.

Equation (5) is the vanilla version of Elastic net. The factor $\lambda_2 \sum_{q=1}^{p} (\beta_q^2)$ allows correlated variables in the corresponding models, which is drawn from ridge regression. If $\alpha$ is equal to 0, the Elastic net keeps the $l_2$-penalty in the model in equation (6) (ridge regression). Similarly, if $\alpha$ is equal to 1, the $l_1$-penalty will only be kept in the Elastic net and equation (6) reduces to a simple LASSO estimator. The $l_2$-norm constraint ensures a unique global minimum in the strictly convex loss function. As in the OP with LASSO estimator, the AIC-type and BIC-type
tuning parameter selectors are employed to the task of selecting the model with the minimum value (see on-line Appendix C). As before, all predictors are standardized before applying the Elastic net estimator.

4.4 CR with LASSO

The CR, originally proposed by Fienberg (1980), was designed for ordinal outcomes in which the categories represent the progression of events or stages in some process. It estimates the probability of one particular category given the categories preceding or succeeding this one. More specifically, it is centred to the binary choice on each ordinal category, which provides the conditional probability of estimating categories. Fienberg (1980), Hardin and Hilbe (2007) and Long and Freese (2006) argue that the CR is superior compared to the binary logistic regression. It is applicable in multi-classification problems where an individual can jump to the discrete rating category without having to pass the intermediate rating categories. Similar to the binary logistic regression, the CR creates binary choices on each ordinal category which allow the calculation of the relevant conditional probabilities. The conditional probability that an individual falls in a level, given that this individual has been beyond that level, is based on “conditional incremental thresholds”. The CR may be regarded as an advanced version of the proportional odds model (ordered logistical model), which preserves the parsimony of the cumulative odds model and considers the ordinal categories of MIRs. These assigned integer values of categories in the CR can be controlled by users, implying that the estimated coefficients in the CR are influenced by the direction chosen for modelling the response variable. In our work, the backward formulation of CR of Archer and Williams (2012) is applied. The progression through the levels of MIRs from investment grade quality (AAA-BBB) to sub-investment grade quality (BB-Below CCC) is expressed by increasing integer values. This helps estimate the odds of lower MIRs rating compared with higher MIRs rating.

The above can be expressed as follows:

\[
\logit(P_{y_{it} = m | y_{it} \leq m, X = A}) = \log \left( \frac{P_{y_{it} = m | y_{it} \leq m, X = A}}{P_{y_{it} < m | y_{it} \leq m, X = A}} \right)
\]

\[13\] Market implied ratings share this property.
\[ A = X_{it}, X_{it-1}, X_{it-2}, X_{it-3}, W_{it}, W_{it-1}, W_{it-2}, W_{it-3}, Z_{it}, Z_{it-1}, Z_{it-2}, Z_{it-3}, y_{it-1}. \] In equation (7), the dependent variable \( y_{it} \) belongs to one of the ordinal rating categories \( m \) described in equation (2). For each unit observation, rather than modeling the response \( y_{it} \) directly, each variable \( y_{it} \) is equal to 1 if the response falls in category \( m \), and 0 otherwise. Thus, conditional likelihood is calculated as in multiple logistic regressions. The above equation can be transformed into the following version to derive the conditional probability:

\[
\Pr(y_{it} = m | y_{it} \leq m) = \frac{e^a}{1 + e^a}, \tag{8}
\]

where \( a = X_{it}\beta_1 + X_{it-1}\beta_2 + X_{it-2}\beta_3 + X_{it-3}\beta_4 + W_{it}\beta_5 + W_{it-1}\beta_6 + W_{it-2}\beta_7 + W_{it-3}\beta_9 + Z_{it}\beta_9 + Z_{it-1}\beta_{10} + Z_{it-2}\beta_{11} + Z_{it-3}\beta_{12} + y_{it-1}\beta_{13}. \)

The parameters can be estimated with maximum likelihood. Similar to OP with LASSO, the maximization of the likelihood proceeds subject to the constraint \( \sum_{q=1}^{p} |\beta_q| \leq s \), where \( s \) is a user-specified tuning parameter. This algorithm combined with LASSO can produce shrinkage coefficients that improve the model’s predictive ability. The resultant model, a CR model, in which a \( l_1 \)-penalized constraint is added into the corresponding likelihood function, is the \( l_1 \)-penalized continuation ratio model. The estimation is presented in equation (9) below:

\[
\hat{\beta} = \arg\max_{\beta} \left( L(\beta | y_{it}, A) - \lambda \sum_{q=1}^{p} |\beta_q| \right), \tag{9}
\]

where \( L(\beta | y_{it}, A) \) is the likelihood function of CR and \( A \) contains the pool of the potential predictors \( (X_{it}, X_{it-1}, X_{it-2}, X_{it-3}, W_{it}, W_{it-1}, W_{it-2}, W_{it-3}, Z_{it}, Z_{it-1}, Z_{it-2}, Z_{it-3}, y_{it-1}). \) All explanatory variables are standardized before applying the LASSO estimator. In line with the abovementioned LASSO counterparts, the AIC-type and the BIC-type tuning parameter
selectors assist with selecting the best model. All predictors are standardized before applying the LASSO estimator.

4.5 CR with Elastic net

To achieve CR with Elastic net, both $l_1$-penalty and $l_2$-penalty are added into maximum likelihood of CR to obtain estimated coefficients, are explained by equations (10) and (11) below:

$$
\hat{\beta} = \arg\max_{\beta} \left( \mathcal{L}(\beta|y_{it}, A) - \lambda_1 \sum_{q=1}^{p} |\beta_q| - \lambda_2 \sum_{q=1}^{p} (\beta_q^2) \right). \quad (10)
$$

Similar to the OP with Elastic net, equation (10) is converted as follows:

$$
\hat{\beta} = \arg\max_{\beta} \left( \mathcal{L}(\beta|y_{it}, A) - \lambda \sum_{q=1}^{p} \left( \alpha |\beta_q| + \frac{1}{2} (1 - \alpha) (\beta_q^2) \right) \right). \quad (11)
$$

where $0 \leq \alpha \leq 1$, $\mathcal{L}(\beta|y_{it}, A)$ is the likelihood function of CR and $A$ contains the pool of the potential predictors ($X_{it}$, $X_{it-1}$, $X_{it-2}$, $X_{it-3}$, $W_{it}$, $W_{it-1}$, $W_{it-2}$, $Z_{it}$, $Z_{it-1}$, $Z_{it-2}$, $Z_{it-3}$, $W_{it}$). All explanatory variables are standardized before applying the Elastic net estimator. Similar to the aforementioned LASSO counterparts, the AIC-type and the BIC-type tuning parameter selectors assist with selecting the best model. Once again, before implementing the Elastic net estimator, all predictors are standardized.

5. Empirical results

5.1 Accuracy

In Table 6 we evaluate the forecasts of the models under study for firms’ CDSIRs and EQIRs, using Accuracy Ratios (ARs). The AR related to the percentage of correct forecast is the sum

---

14 For more details, see on-line Appendix C.
of all diagonal terms divided by the total number of observations in each contingency table (see Appendix B) \( AR = \frac{1}{T} \sum_{t=1}^{T} 1(\hat{q}_t = q_t) \) where \( \hat{q}_t \) denotes the predicted rating and \( q_t \) represents the actual outcome. We report statistics for all candidate models for both in- and out-of-sample predictions. For the out-of-sample predictions of ratings we use the past and current information available up to time \( T \). We use an expanding window method, which allows the successive observations to be included in the initial sample prior to forecast of the next one-step ahead prediction of the rating while keeping the start date of the sample fixed. By this method, we forecast future ratings \( \hat{q}_{t+1}, \hat{q}_{t+2} \) etc. The initial estimation window is 2002 to 2005 and the first prediction date is year 2006. We then increase \( T \) by one each time until \( T \) reaches year 2008. In addition, we report at the foot of each panel the number of surviving variables\(^{15}\).

To begin with the in-sample exercise, we find no notable differences between the competing models since they present a similar in-sample performance for both types of market implied ratings. With respect to CDSIRs, approximately 90% predictions are correct, while for EQIRs we find that the models have approximately 95% correct predictions. Moving to the out-of-sample prediction, the results suggest that the LASSO models clearly outperform their OP benchmark. When considering the CDSIRs, our results indicate that the percentage of correct predictions increase from 22% in the OP model to 84% in the LASSO models. With reference to EQIRs, the percentage of correct predictions improves from 49% in the OP model to 91% in the LASSO models.

Next, we compare the within performance of the LASSO models, by considering alternative LASSO specifications. Starting with CDSIRs, the models with BIC-type tuning parameter selector provide more accurate out-of-sample forecasts compared to the LASSO models with AIC-type tuning parameter selector. For EQIRs, there is little difference in the accuracy ratios of the various LASSO candidate models. It is interesting to note that the LASSO models with BIC-type tuning parameter selector select consistently a smaller number of predictors than

\(^{15}\) The surviving variables are defined as predictors with non-zero estimated coefficients after the penalized procedure. In the benchmark model, we do not drop any variables since OP does not penalise regression coefficients.
their AIC-type counterparts. This does not seem to affect their forecasting performance for EQIRs but leads to more accurate predictions for CDSIRs.\textsuperscript{16}

5.2 Statistical significance

To evaluate the relative performance of the models presented in the sub-section above, we employ the Stuart–Maxwell test. This approach will help us formally test for the statistical significance of the difference between forecasts and to further validate our main findings. The Stuart–Maxwell test (Stuart, 1955; Maxwell, 1970) is a generalized version of McNemar’ test (McNemar, 1947), which is associated with multiple \((k)\) categories and tests whether the difference between two related samples from an ordinal field is statistically different from zero. The Stuart–Maxwell tests the null hypothesis of equal marginal proportion for each category between the forecasts of two models (model A vs model B). Under the null hypothesis, the statistic is distributed as chi-square with \(k-1\) degrees of freedom. A statistically significant Stuart–Maxwell test statistic indicates that the forecasts of the first model (A) are different from those of the second model (B).

For the CDSIRs, as can be seen from Table 6, there is evidence of a statistically significant difference in out-of-sample predictions between the OP model and other competing models. For the EQIRs, the results of the out-of-sample forecasts paint a very similar picture. However, this tendency cannot be clearly observed in the in-sample forecasts of all competing models for both CDSIRs and EQIRs. Combing these statistics with the accuracy ratios, it can be confirmed that adding LASSO or the Elastic net estimator in the OP model or the CR model, can produce different out-of-sample forecasts that are more accurate than those generated by the OP.

5.3 Robustness tests

The findings of the previous Section are further validated by carrying-out several robustness tests. In the first test, we consider an alternative choice of tuning parameters by using the

\textsuperscript{16} Tables B1 to B36 in on-line Appendix B illustrate the contingency tables of the predicted against the actual outcome both in- and out-of-sample results for the various models presented in Table 6.
cross validation. Next, we rely on an alternative benchmark for forecasting, namely the Principal Component Analysis. In a further test we employ a tuning-free version of the LASSO estimator. As a fourth test, we use alternative measures of predictive ability. Furthermore, we allow only investment grade ratings. Finally, we consider random effects in panel data.

5.3.1 Cross validation

It is well accepted that the choice of tuning parameter is crucial for the finite-sample performance of estimators such as LASSO or elastic net. Motivated by this consideration, we present forecasts when cross validation is employed, which is one of the most commonly used model selection criteria (see for instance, Stone, 1974 and Yang, 2007). Specifically, we use the ten-fold stratified cross validation (10-fold CV) as a general application of tuning parameter selector. This choice is in line with the relevant literature of model selection (see for example, Kohavi, 1995).

Insert Table 7

The Accuracy Ratios and Stuart–Maxwell test statistics, presented in Table 7, corroborate our main findings. In the in-sample forecast exercises, all models present similar predictive ability. In the out-of-sample prediction, however, for the CDSIRs, all competing LASSO models provide more accurate forecasts than the benchmark model. This tendency can be also observed in the EQIRs. In sum, it appears that there are significant gains in predictive ability once the LASSO is applied even when cross validation is employed. We conclude that our main results are robust to alternative tuning choices.

5.3.2 Principal Component Analysis with OP

One of the most popular statistical procedures for variable selection is the Principal Component Analysis (PCA). PCA converts a set of possibly correlated variables to a smaller set of uncorrelated variables called Principal Components (PC). The first PC accounts as much of the variability in the dataset as possible and each succeeding component turn attains the highest variance possible under the constraint that it is orthogonal to the preceding
components. PCA is probably the most popular dimension reduction procedure in Economics and Finance and has been applied successfully to a series of forecasting problems (see Stock and Watson, 2002, Ludvigson and Ng, 2009 and Bailey et al, 2016). In our application, we are dealing with a large set of possibly correlated variables and thus a natural candidate to benchmark our procedure is the PCA. We select the PCs that account for the 70%, 80% and 90% of the variability in our dataset and combine them with OP. The relevant accuracy ratios are presented in Table 8.

Comparing Tables 8 and 6 (where the accuracy ratios of OP and OP with LASSO are presented), we note that PCA naturally improves the out-of-sample accuracy of the single OP model. For our OP combined with LASSO models, the only case that PCA manages to beat our models in terms of out-of-sample accuracy is for CDSIRs. In that case, OP based on the PCs that explain the 70% and 90% of the variability, present a higher accuracy ratio compared to LASSO model with AIC-type tuning parameter selector (but not under the BIC-type selector). It is also interesting to note that PCA with OP displays lower in-sample accuracy ratios at all cases. These results allow to argue that our LASSO formulations are robust to an alternative benchmark.

5.3.3 A tuning-free version of the LASSO

While the results presented so far are robust to different tuning choices, including cross validation, it is important to note that the latter is computationally costly and theoretically less well developed, especially for the purpose of variable selection and the estimation of regression coefficients (see Sun and Zhang, 2012). Thus, to further alleviate potential concerns regarding the choice of the tuning parameter, we employ the scaled LASSO, developed by Sun and Zhang (2012), without depending on model selection criteria such as AIC, BIC or CV\textsuperscript{17,18}.

\textit{Insert Table 9}

\textsuperscript{17} Another tuning-free version of the LASSO estimator is the square-root LASSO by Belloni et al. (2011).

\textsuperscript{18} For the scaled LASSO, the authors generated the gradient descent algorithm in a convex minimization of a penalized joint loss function.
Table 9 presents the relevant Accuracy Ratios and the realizations of the Stuart–Maxwell statistical tests. We note that on in-sample evidence all models continue to display similar accuracy. In the out-of-sample evaluation, for the CDSIRs and EQIRs the predictive performance of the scaled LASSO is superior compared with that of the OP. Once again, this finding is consistent with our main results, indicating that our findings are robust to using the scaled LASSO.

5.3.4 Alternative measures of predictive ability

5.3.4.1 Thus far, the relative performance of the estimated models is evaluated in terms of an informal goodness of fit indicator, by comparing predicted and observed ratings. It is possible, however, to give a more quantitative measure of the predictive ability of our models. We therefore check the robustness of our measure of predictive power by using a measure based on a technique proposed by Merton (1981) and used in Henriksson and Merton (1981), Pesaran and Timmermann (1994), Kim et al. (2008) and Mizen and Tsoukas (2012). Specifically, let $CP_j$ be the proportion of the correct predictions made by $\hat{q}_t$ when the true state is given by $q_t = j$. From the definition of conditional probability, $CP$ is computed as $CP_j = \frac{\sum_{t=1}^{T} I(q_t = j|q_t = j)}{\sum_{t=1}^{T} I(q_t = j)}$, and the Merton’s correct measure expressed $CP$ is given by $CP = \frac{1}{J-1}[\sum_{j=1}^{J} CP_j - 1]$ where $J$ is the number of categories, and $-\frac{1}{J-1} < CP < 1$. In the contingency table (see Appendix B) $CP$ is the unweighted average of $CP_j$’s minus one (to correct for the phenomenon that certain categories are over-represented). The $CP_j$’s are calculated as the proportion of correct predictions divided by the total of each row. This modifies the measure of predictive ability to discount the influence of the dominant outcome. A high $CP$ score indicates that the predictor is accurate for all rating categories.

Insert Table 10

The Accuracy Ratios when we account for the influence of the dominant outcome by reporting the Merton correct predictions are shown in Table 10. We also report the corresponding Stuart–Maxwell statistical tests. The test produces $CP$ ratios that confirm our
main findings. In the in-sample exercise, there is little difference between the OP and the LASSO models. In contrast, the predictive ability of the out-of-sample predictions is superior when penalty functions are applied.

5.3.4.2
An alternative measure for our forecasts can be constructed based on the misclassification rate of Hastie et al. (2009): $T^{-1} \sum_{t=1}^{T} I(\hat{q}_t \neq q_t)$. In the spirit of Diebold and Mariano (1995), we estimate the misclassification rate for each forecast and then we test if the mean difference of these rates between two models is 0 or not. If this difference is statistically different than 0, then this means that the two models generate different forecasts. Tables 11 and 12 present the relevant p-values for our out-of-sample forecasts.

Insert Tables 11
Insert Tables 12

We note that in almost all cases, our forecasts are statistically different. These results supplement the previous section and further demonstrate the superiority of LASSO as a variable selection technique and the effectiveness of BIC criterion in tuning the LASSO parameters. To sum up, the results are robust to using a test that calculates correct predictions using the proportion of correct predictions for each of the various rating categories.

5.3.5 Investment grade ratings

Much of the previous related literature studies employ data with investment grade ratings. However, as noted by Amato and Furfine (2004), restricting attention to one category is likely to induce selection bias. On the other hand, pooling together both categories may result in misspecification of our model if changes in financial and business risk have a different impact on creditworthiness across the groups of firms. Therefore, we drop all speculative grade ratings and re-estimate our models.

Insert Table 13
The results in Table 13 corroborate our main findings. In the in-sample forecast exercises, all models present similar performance. For the CDSIRs, both versions of the CR with LASSO provide more accurate forecasts. On the other hand, for EQIRs all LASSO models display better predictive performance than their OP benchmark. To sum up, even when limiting our sample to investment grade ratings only, the out-of-sample predictions evidence that the LASSO models outperform the benchmark model.

5.3.6 Accounting for the panel data dimension

As a final robustness test we consider a random-effects version of the ordered probit model to take into account the panel data dimension of the data-set. The Accuracy Ratios and the corresponding Stuart–Maxwell statistical tests are reported in Table 14. As can be seen, the main findings remain unchanged: when we apply the LASSO or Elastic net estimator the models have superior predictive ability compared to the ordered probit model, even when random effects are included. We conclude that our findings are robust to estimating the models with random effects to deal with the panel data nature of the sample. 

Insert Table 14

5.4 Discussion

In the previous Sections, a forecasting exercise on CDSIRs and EQIRs prediction is presented. For both types of market implied ratings all models present similar in-sample accuracy. In the out-of-sample evaluation, for CDSIRs, we observe the LASSO models controlled by BIC-type tuning parameter selector outperform their benchmarks. We also observe that these models, select a smaller number of surviving variables than their counterparts with the AIC-type selector. This lends support to the argument that the models with BIC-type tuning parameter selector make better use of the available information. For the EQIRs, we note a similar pattern. The models with the BIC-type tuning parameter selector outperform their counterparts with the AIC-type selector in terms of accuracy while at the same time they select a smaller number of variables.
To sum up, we note that the LASSO models are able to provide more accurate out-of-sample forecasts on the CDSIRs and EQIRs ratings that outperform the OP model. This is of particular interest given that the OP model dominates the related literature in predicting credit ratings. From the LASSO models under study, the optimized models with BIC-type tuning parameter selector seem able to provide better forecasts while at the same time they use less predictors. These results are robust to modifying the tuning parameters, to considering a tuning-free version of LASSO, to evaluating the predictive performance of the models using a different statistical measure and to restricting the dataset to investment grade ratings.

6 Conclusion

The ability to predict credit ratings within a reasonable margin of accuracy is of vital importance for both market participants and rating agencies. The focus on market implied ratings is even more justified as long-term ratings have been heavily criticized about their performance during the recent global financial crisis. We model the prediction of market implied ratings applying a variable selection technique, the least absolute shrinkage and selection operator (LASSO), and its most promising derivation, the Elastic net, into ordered probit and continuation ratio models. All LASSO models select the most relevant predictors from a set of 268 variables and forecast the MIRs for a period of six years (2002 to 2008). This marks a break with the existing literature which typically relies on discrete limited dependent variable models.

Our results using monthly data from the US offer several interesting results. First, we show that financial factors along with market-driven and macroeconomic variables contain information about market implied ratings. Second, the LASSO models perform better in out-of-sample prediction than do ordered probit models, mostly adopted in previous studies. Finally, the optimized LASSO models with BIC-type tuning parameter selector outperform their counterparts with AIC-type selector for the dataset and periods under study. Hence, LASSO-selected models attain an improved forecasting power.
These results should go further in convincing risk managers and academics to explore variable selection models when assessing credit risk. The structure of credit ratings is unknown and likely to vary through time. Limited dependent variable models require a-priori knowledge on the explanatory variables set, which can lead to misspecifications. On the other hand, variable selection models such as LASSO, are more flexible and can unveil the underlying structure of the problem leading to superior estimations and improved predictive ability.

References


<table>
<thead>
<tr>
<th>Table 1: Rating categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Implied Ratings</td>
</tr>
<tr>
<td>AAA</td>
</tr>
<tr>
<td>AA</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>BBB</td>
</tr>
<tr>
<td>BB</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>Below CCC</td>
</tr>
</tbody>
</table>

Notes: The table presents the rating categories and the corresponding numerical values.

<table>
<thead>
<tr>
<th>Table 2: CDSIRs of firms by year</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>2002</td>
</tr>
<tr>
<td>2003</td>
</tr>
<tr>
<td>2004</td>
</tr>
<tr>
<td>2005</td>
</tr>
<tr>
<td>2006</td>
</tr>
<tr>
<td>2007</td>
</tr>
<tr>
<td>2008</td>
</tr>
</tbody>
</table>

Number of Observations: 69 242 350 397 183 48 9 1298

Notes: The table presents the distribution of firms by rating category for CDSIRs by year.

<table>
<thead>
<tr>
<th>Table 3: EQIRs of firms by year</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>2002</td>
</tr>
<tr>
<td>2003</td>
</tr>
<tr>
<td>2004</td>
</tr>
<tr>
<td>2005</td>
</tr>
<tr>
<td>2006</td>
</tr>
<tr>
<td>2007</td>
</tr>
<tr>
<td>2008</td>
</tr>
</tbody>
</table>

Number of Observations: 5 97 516 586 292 44 0 1540

Notes: The table presents the distribution of firms by rating category for EQIRs by year.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean (1)</th>
<th>Standard Deviation (2)</th>
<th>Minimum (3)</th>
<th>Maximum (4)</th>
<th>p-value (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DETAy</td>
<td>9.6388</td>
<td>1.0132</td>
<td>7.2272</td>
<td>12.2084</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>8.9536</td>
<td>0.9959</td>
<td>6.2539</td>
<td>12.2087</td>
<td>0.0000</td>
</tr>
<tr>
<td>AE</td>
<td>3.2256</td>
<td>3.6336</td>
<td>1.3324</td>
<td>73.7340</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>5.0224</td>
<td>8.8729</td>
<td>1.3283</td>
<td>123.5602</td>
<td>0.0000</td>
</tr>
<tr>
<td>LDA</td>
<td>20.2416</td>
<td>11.2733</td>
<td>0.0000</td>
<td>79.3983</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>30.1868</td>
<td>19.4641</td>
<td>0.0000</td>
<td>110.4453</td>
<td>0.0000</td>
</tr>
<tr>
<td>TDA</td>
<td>24.3546</td>
<td>11.9238</td>
<td>1.3017</td>
<td>87.0595</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>37.4629</td>
<td>20.8746</td>
<td>0.6518</td>
<td>126.8760</td>
<td>0.0000</td>
</tr>
<tr>
<td>TDEBITDA</td>
<td>2.3395</td>
<td>1.2107</td>
<td>0.3200</td>
<td>19.6400</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>4.1208</td>
<td>2.8918</td>
<td>0.3200</td>
<td>23.0700</td>
<td>0.0000</td>
</tr>
<tr>
<td>EBITINT</td>
<td>13.3805</td>
<td>19.0270</td>
<td>0.1541</td>
<td>209.3023</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>7.2685</td>
<td>15.2730</td>
<td>0.1248</td>
<td>210.4054</td>
<td>0.0000</td>
</tr>
<tr>
<td>EBITDAINT</td>
<td>16.0444</td>
<td>21.9950</td>
<td>0.6500</td>
<td>235.4000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>9.0872</td>
<td>14.5213</td>
<td>0.3100</td>
<td>196.7000</td>
<td>0.0000</td>
</tr>
<tr>
<td>CFOA</td>
<td>6.9186</td>
<td>6.2625</td>
<td>-41.0623</td>
<td>38.4372</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>7.1077</td>
<td>7.2097</td>
<td>-13.4106</td>
<td>65.9955</td>
<td>0.0000</td>
</tr>
<tr>
<td>CASHEQA</td>
<td>8.6116</td>
<td>9.6531</td>
<td>0.0238</td>
<td>71.8277</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>6.5258</td>
<td>8.1009</td>
<td>0.0030</td>
<td>64.0272</td>
<td>0.0000</td>
</tr>
<tr>
<td>OM</td>
<td>13.9496</td>
<td>9.7304</td>
<td>-33.1555</td>
<td>53.0189</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>10.3893</td>
<td>9.9617</td>
<td>-20.2276</td>
<td>52.6046</td>
<td>0.0000</td>
</tr>
<tr>
<td>ROC</td>
<td>3.7919</td>
<td>5.3598</td>
<td>-34.1330</td>
<td>35.7771</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>2.1389</td>
<td>7.0866</td>
<td>-36.3880</td>
<td>34.5209</td>
<td>0.0000</td>
</tr>
<tr>
<td>ROE</td>
<td>12.5162</td>
<td>35.3349</td>
<td>-361.1511</td>
<td>452.2565</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
Notes: The Table reports the summary statistics of the explanatory variables used in the empirical models. Column 5 reports the p-value for the test of equality of means between the investment grade and non-investment grade categories. Investment grade refers to ratings from AAA to BBB. Non-investment grade refers to ratings BB and below. A detailed description of the variables used in this study is given in Table A1 in the online Appendix A.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Non-investment grade</th>
<th>Investment grade</th>
<th>Non-investment grade</th>
<th>Investment grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROA</td>
<td>9.5597</td>
<td>3.9430</td>
<td>2.5165</td>
<td>40.7412</td>
</tr>
<tr>
<td>FFD</td>
<td>46.9441</td>
<td>4.4180</td>
<td>5.5633</td>
<td>32.5483</td>
</tr>
<tr>
<td>LIQ</td>
<td>-359.7868</td>
<td>-26.4074</td>
<td>-23.8888</td>
<td>-16.2800</td>
</tr>
<tr>
<td>EXRET</td>
<td>516.7883</td>
<td>22.0518</td>
<td>23.2336</td>
<td>267.1700</td>
</tr>
<tr>
<td>STD</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>BETA</td>
<td>0.0171</td>
<td>0.0195</td>
<td>0.0117</td>
<td>0.0195</td>
</tr>
<tr>
<td>PD1</td>
<td>0.0646</td>
<td>0.1108</td>
<td>0.0646</td>
<td>0.1108</td>
</tr>
<tr>
<td>PD5</td>
<td>-0.3525</td>
<td>-0.3526</td>
<td>0.4673</td>
<td>0.4673</td>
</tr>
<tr>
<td>Beta</td>
<td>0.4527</td>
<td>0.4673</td>
<td>0.4527</td>
<td>0.4673</td>
</tr>
<tr>
<td>Beta</td>
<td>0.0149**</td>
<td>0.0149**</td>
<td>0.0149**</td>
<td>0.0149**</td>
</tr>
</tbody>
</table>

Non-investment grade refers to ratings BB and below. Investment grade refers to ratings from AAA to BBB.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean (1)</th>
<th>Standard Deviation (2)</th>
<th>Minimum (3)</th>
<th>Maximum (4)</th>
<th>p-value (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DETA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>9.5314</td>
<td>1.0515</td>
<td>7.0031</td>
<td>12.2084</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>8.8701</td>
<td>0.9461</td>
<td>6.2539</td>
<td>12.2087</td>
<td></td>
</tr>
<tr>
<td>AE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>3.4552</td>
<td>4.7315</td>
<td>1.3283</td>
<td>116.1204</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>5.3173</td>
<td>9.3934</td>
<td>1.3283</td>
<td>123.5602</td>
<td></td>
</tr>
<tr>
<td>LDA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>20.9162</td>
<td>12.5949</td>
<td>0.0000</td>
<td>110.4453</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>32.5301</td>
<td>19.9743</td>
<td>0.0000</td>
<td>110.1548</td>
<td></td>
</tr>
<tr>
<td>SDA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>3.7397</td>
<td>4.1512</td>
<td>0.3200</td>
<td>23.4635</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>3.0151</td>
<td>3.9714</td>
<td>0.3200</td>
<td>23.5410</td>
<td></td>
</tr>
<tr>
<td>TDA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>25.3924</td>
<td>13.5320</td>
<td>1.3017</td>
<td>126.8209</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>40.3718</td>
<td>21.1606</td>
<td>0.6518</td>
<td>126.8760</td>
<td></td>
</tr>
<tr>
<td>TDEBITDA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>2.5062</td>
<td>1.4809</td>
<td>0.3200</td>
<td>20.4000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>4.4812</td>
<td>3.0369</td>
<td>0.3200</td>
<td>23.0700</td>
<td></td>
</tr>
<tr>
<td>EBITINT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>13.2588</td>
<td>20.3637</td>
<td>0.1601</td>
<td>210.4054</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>5.3897</td>
<td>10.2101</td>
<td>0.1248</td>
<td>157.3792</td>
<td></td>
</tr>
<tr>
<td>EBITDAINT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>15.7294</td>
<td>21.3681</td>
<td>0.4100</td>
<td>235.4000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>7.2437</td>
<td>11.9978</td>
<td>0.3100</td>
<td>171.8000</td>
<td></td>
</tr>
<tr>
<td>CASHEQA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>7.8973</td>
<td>6.5051</td>
<td>-41.0623</td>
<td>40.5546</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>5.9413</td>
<td>7.0742</td>
<td>-41.0623</td>
<td>65.9955</td>
<td></td>
</tr>
<tr>
<td>EBITA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>8.6834</td>
<td>9.6845</td>
<td>0.0030</td>
<td>71.8277</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>5.7724</td>
<td>7.3325</td>
<td>0.0033</td>
<td>58.4741</td>
<td></td>
</tr>
<tr>
<td>OM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>13.0445</td>
<td>9.5560</td>
<td>-16.9133</td>
<td>53.0189</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>10.3826</td>
<td>10.3917</td>
<td>-20.2276</td>
<td>52.5504</td>
<td></td>
</tr>
<tr>
<td>ROC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>3.8570</td>
<td>5.6971</td>
<td>-35.0754</td>
<td>35.7771</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>1.5077</td>
<td>7.1322</td>
<td>-36.3880</td>
<td>34.5140</td>
<td></td>
</tr>
<tr>
<td>EBITA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>13.3395</td>
<td>36.3731</td>
<td>-361.1511</td>
<td>473.0769</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>7.5924</td>
<td>48.9666</td>
<td>-359.7868</td>
<td>516.7883</td>
<td></td>
</tr>
<tr>
<td>ROA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>4.1096</td>
<td>4.6843</td>
<td>-26.4074</td>
<td>23.2190</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>1.8536</td>
<td>5.4585</td>
<td>-26.4074</td>
<td>23.2336</td>
<td></td>
</tr>
<tr>
<td>FFDA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>38.8160</td>
<td>31.7571</td>
<td>-16.2800</td>
<td>267.1700</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>21.6264</td>
<td>24.5682</td>
<td>-17.8000</td>
<td>224.0300</td>
<td></td>
</tr>
<tr>
<td>LIQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>13.5432</td>
<td>11.8937</td>
<td>-13.5675</td>
<td>62.1817</td>
<td>0.0000</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>9.495881</td>
<td>11.69661</td>
<td>-13.56745</td>
<td>57.2284</td>
<td></td>
</tr>
<tr>
<td>EXRET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>0.0140</td>
<td>0.0723</td>
<td>-0.3525</td>
<td>0.4648</td>
<td>0.5831</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>0.0181</td>
<td>0.1089</td>
<td>-0.3526</td>
<td>0.4673</td>
<td></td>
</tr>
</tbody>
</table>
The Table reports the summary statistics of the explanatory variables used in the empirical models. Column 5 reports the p-value for the test of equality of means between the investment grade and non-investment grade categories. Investment grade refers to ratings from AAA to BBB. Non-investment grade refers to ratings BB and below. A detailed description of the variables used in this study is given in Table A1 in the online Appendix A.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Investment grade</th>
<th>Non-investment grade</th>
<th>p-value</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSIZE</td>
<td>0.2047</td>
<td>0.0922</td>
<td>0.0088</td>
<td>1.7570</td>
<td>0.0000</td>
</tr>
<tr>
<td>STD</td>
<td>0.0172</td>
<td>0.0250</td>
<td>0.0041</td>
<td>0.1134</td>
<td>0.0000</td>
</tr>
<tr>
<td>BETA</td>
<td>0.9081</td>
<td>1.1028</td>
<td>-0.8757</td>
<td>4.8848</td>
<td>0.0000</td>
</tr>
<tr>
<td>PD1</td>
<td>24.1369</td>
<td>215.5352</td>
<td>2.0000</td>
<td>3000.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>PD5</td>
<td>250.8945</td>
<td>1010.6400</td>
<td>14.0000</td>
<td>4495.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Notes: The Table reports the summary statistics of the explanatory variables used in the empirical models. Column 5 reports the p-value for the test of equality of means between the investment grade and non-investment grade categories. Investment grade refers to ratings from AAA to BBB. Non-investment grade refers to ratings BB and below. A detailed description of the variables used in this study is given in Table A1 in the online Appendix A.
<table>
<thead>
<tr>
<th>Model Type</th>
<th>OP</th>
<th>OP_LASSO</th>
<th>OP_ELASTIC NET</th>
<th>CR_LASSO</th>
<th>CR_ELASTIC NET</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AIC</td>
<td>BIC</td>
<td>AIC</td>
<td>BIC</td>
<td>AIC</td>
</tr>
<tr>
<td><strong>CDSIRs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-sample prediction</td>
<td>90.23%</td>
<td>90.30%*</td>
<td>89.39%***</td>
<td>90.30%*</td>
<td>89.80%***</td>
</tr>
<tr>
<td>Out-of-sample prediction</td>
<td>22.02%</td>
<td>31.56%***</td>
<td>84.53%***</td>
<td>31.56%***</td>
<td>83.73%***</td>
</tr>
<tr>
<td>Surviving variables</td>
<td>268</td>
<td>143</td>
<td>45</td>
<td>144</td>
<td>45</td>
</tr>
<tr>
<td><strong>EQIRs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-sample prediction</td>
<td>95.05%</td>
<td>95.00%</td>
<td>94.69%*</td>
<td>95.05%</td>
<td>94.52%*</td>
</tr>
<tr>
<td>Out-of-sample prediction</td>
<td>48.98%</td>
<td>84.77%***</td>
<td>90.95%***</td>
<td>80.88%***</td>
<td>90.78%***</td>
</tr>
<tr>
<td>Surviving variables</td>
<td>268</td>
<td>167</td>
<td>95</td>
<td>181</td>
<td>87</td>
</tr>
</tbody>
</table>

Notes: The Table reports the Accuracy Ratios and the number of surviving variables for each model under study. "OP" stands for the ordered probit model. "OP_LASSO" refers to the ordered probit model with LASSO estimator. "OP_ELASTIC NET" stands for the ordered probit model with Elastic net estimator. "CR_LASSO" indicates the continuation ratio model with LASSO estimator. "CR_ELASTIC NET" is the continuation ratio model with Elastic net estimator. "AIC" is the AIC-type tuning parameter selector. "BIC" is the BIC-type tuning parameter selector. *** denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 1% significance level. * denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 5% significance level. ** denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 10% significance level.
### Table 8: PCA combined with OP

<table>
<thead>
<tr>
<th></th>
<th>PCA with OP</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>CDSIRs</td>
<td>In-sample prediction</td>
<td>45.94%</td>
<td>51.36%</td>
</tr>
<tr>
<td></td>
<td>Out-of-sample prediction</td>
<td>45.71%</td>
<td>32.01%</td>
</tr>
<tr>
<td></td>
<td>Selected principal components</td>
<td>16</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>PCs cumulative percentage of total variation</td>
<td>70.92%</td>
<td>80.00%</td>
</tr>
<tr>
<td>EQIRs</td>
<td>In-sample prediction</td>
<td>63.61%</td>
<td>72.00%</td>
</tr>
<tr>
<td></td>
<td>Out-of-sample prediction</td>
<td>63.37%</td>
<td>54.74%</td>
</tr>
<tr>
<td></td>
<td>Selected principal components</td>
<td>15</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>PCs cumulative percentage of total variation</td>
<td>71.23%</td>
<td>80.43%</td>
</tr>
</tbody>
</table>

Notes: The Table reports the Accuracy Ratios and the number of PCs. "OP" stands for the ordered probit model, "PC" for principal components and "PCA" for principal component analysis.

### Table 9: Accuracy Ratios and selected variables (scaled LASSO)

<table>
<thead>
<tr>
<th></th>
<th>OP</th>
<th>SCALED_LASSO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In-sample prediction</td>
<td>90.23%</td>
</tr>
<tr>
<td>CDSIRs</td>
<td>Out-of-sample prediction</td>
<td>22.02%</td>
</tr>
<tr>
<td></td>
<td>Surviving variables</td>
<td>268</td>
</tr>
<tr>
<td></td>
<td>In-sample prediction</td>
<td>95.05%</td>
</tr>
<tr>
<td>EQIRs</td>
<td>Out-of-sample prediction</td>
<td>48.98%</td>
</tr>
<tr>
<td></td>
<td>Surviving variables</td>
<td>268</td>
</tr>
</tbody>
</table>

Notes: The Table reports the Accuracy Ratios and the number of surviving variables for each model under study. "OP" stands for the ordered probit model. "SCALED_LASSO" refers to the scaled LASSO for tuning-free parameter. *** denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 1% significance level. ** denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 5% significance level. * denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 10% significance level.
Table 10: CP Ratios and selected variables

<table>
<thead>
<tr>
<th></th>
<th>OP</th>
<th>OP_LASSO</th>
<th>OP_ELASTIC NET</th>
<th>CR_LASSO</th>
<th>CR_ELASTIC NET</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AIC</td>
<td>BIC</td>
<td>AIC</td>
<td>BIC</td>
<td>AIC</td>
</tr>
<tr>
<td>CDSIRs</td>
<td>In-sample prediction</td>
<td>69.29%</td>
<td>66.89%*</td>
<td>61.04%***</td>
<td>66.95%</td>
</tr>
<tr>
<td></td>
<td>Out-of-sample prediction</td>
<td>18.22%</td>
<td>25.49%***</td>
<td>61.01%***</td>
<td>25.49%***</td>
</tr>
<tr>
<td></td>
<td>Surviving variables</td>
<td>268</td>
<td>143</td>
<td>45</td>
<td>144</td>
</tr>
<tr>
<td>EQIRs</td>
<td>In-sample prediction</td>
<td>78.30%</td>
<td>77.16%</td>
<td>74.46%*</td>
<td>77.19%</td>
</tr>
<tr>
<td></td>
<td>Out-of-sample prediction</td>
<td>35.49%</td>
<td>66.50%***</td>
<td>70.10%***</td>
<td>63.42%***</td>
</tr>
<tr>
<td></td>
<td>Surviving variables</td>
<td>268</td>
<td>167</td>
<td>95</td>
<td>181</td>
</tr>
</tbody>
</table>
|        | Notes: The Table reports the CP ratios and the number of surviving variables for each model under study. "OP" stands for the ordered probit model. "OP_LASSO" refers to the ordered probit model with LASSO estimator. "OP_ELASTIC NET" stands for the ordered probit model with Elastic net estimator. "CR_LASSO" indicates the continuation ratio model with LASSO estimator. "CR_ELASTIC NET" is the continuation ratio model with Elastic net estimator. "AIC" is the AIC-type tuning parameter selector. "BIC" is the BIC-type tuning parameter selector. *** denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 1% significance level. ** denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 5% significance level. * denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 10% significance level.
Table 11: Equal performance tests of out-of-sample prediction of CDSIRs

<table>
<thead>
<tr>
<th></th>
<th>OP</th>
<th>OP_LASSO</th>
<th>OP_ELASTIC NET</th>
<th>CR_LASSO</th>
<th>CR_ELASTIC NET</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AIC</td>
<td>BIC</td>
<td>AIC</td>
<td>BIC</td>
<td>AIC</td>
</tr>
<tr>
<td>OP</td>
<td>~</td>
<td>7.60***</td>
<td>38.38***</td>
<td>7.60***</td>
<td>38.38***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32.58***</td>
<td>~</td>
<td>32.58***</td>
<td>~</td>
</tr>
<tr>
<td>OP_LASSO</td>
<td>~</td>
<td>-32.58***</td>
<td>~</td>
<td>-32.58***</td>
<td>~</td>
</tr>
<tr>
<td></td>
<td></td>
<td>~</td>
<td>32.58***</td>
<td>~</td>
<td>32.58***</td>
</tr>
<tr>
<td>OP_ELASTIC NET</td>
<td>~</td>
<td>-17.08***</td>
<td>-1.62</td>
<td>-17.08***</td>
<td>-1.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>~</td>
<td>17.36***</td>
<td>~</td>
<td>17.36***</td>
</tr>
<tr>
<td>CR_LASSO</td>
<td>~</td>
<td>-16.27***</td>
<td>1.73*</td>
<td>~</td>
<td>-16.27***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>~</td>
<td>16.35***</td>
<td>~</td>
<td>16.35***</td>
</tr>
</tbody>
</table>

Note: The table reports the test of equality of the mean difference in losses for two models. *** denotes that the null hypothesis of equal performance of two models is rejected at the 1% significant level. ** denotes that the null hypothesis of equal performance of two models is rejected at the 5% significant level and * denotes that the null hypothesis of equal performance of two models is rejected at the 10% significant level. ~ indicates that the two models generate the same set of forecasts.

Table 12: Equal performance tests of out-of-sample prediction of EQIRs

<table>
<thead>
<tr>
<th></th>
<th>OP</th>
<th>OP_LASSO</th>
<th>OP_ELASTIC NET</th>
<th>CR_LASSO</th>
<th>CR_ELASTIC NET</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AIC</td>
<td>BIC</td>
<td>AIC</td>
<td>BIC</td>
<td>AIC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.03***</td>
<td>-5.83***</td>
<td>5.82***</td>
<td>1.79*</td>
</tr>
<tr>
<td>OP_LASSO</td>
<td>~</td>
<td>-8.79***</td>
<td>-1.41</td>
<td>-5.46***</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.64***</td>
<td>6.20***</td>
<td>8.85***</td>
<td>-0.26</td>
</tr>
<tr>
<td>OP_ELASTIC NET</td>
<td>~</td>
<td>-5.25***</td>
<td>1.73*</td>
<td>-5.25***</td>
<td>-5.81***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.57***</td>
<td>-6.37***</td>
<td>5.46***</td>
<td>5.46***</td>
</tr>
<tr>
<td>CR_LASSO</td>
<td>~</td>
<td>-8.71***</td>
<td>-1.00</td>
<td>-8.71***</td>
<td>-1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>~</td>
<td>8.66***</td>
<td>~</td>
<td>8.66***</td>
</tr>
</tbody>
</table>

Note: The table reports the test of equality of the mean difference in losses for two models. *** denotes that the null hypothesis of equal performance of two models is rejected at the 1% significant level. ** denotes that the null hypothesis of equal performance of two models is rejected at the 5% significant level, * denotes that the null hypothesis of equal performance of two models is rejected at the 10% significant level. ~ indicates that the two models generate the same set of forecasts.
Table 13: Accuracy Ratios and selected variables (investment-grade ratings)

<table>
<thead>
<tr>
<th></th>
<th>OP</th>
<th>OP_LASSO</th>
<th>OP_ELASTIC NET</th>
<th>CR_LASSO</th>
<th>CR_ELASTIC NET</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AIC</td>
<td>BIC</td>
<td>AIC</td>
<td>BIC</td>
<td>AIC</td>
</tr>
<tr>
<td>In-sample prediction</td>
<td>90.76%</td>
<td>90.45%**</td>
<td>90.14%**</td>
<td>89.42%***</td>
<td>90.56%**</td>
</tr>
<tr>
<td>Out-of-sample prediction</td>
<td>49.18%</td>
<td>77.04%***</td>
<td>84.33%***</td>
<td>84.22%***</td>
<td>84.87%***</td>
</tr>
</tbody>
</table>

Surviving variables

<table>
<thead>
<tr>
<th>CDSRs</th>
<th>In-sample prediction</th>
<th>OP_re</th>
<th>OP_re_LASSO</th>
<th>OP_re_ELASTIC NET</th>
<th>CR_re</th>
<th>CR_re_LASSO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In-sample prediction</td>
<td>83.62%</td>
<td>84.15%***</td>
<td>83.00%***</td>
<td>84.03%**</td>
<td>83.00%***</td>
</tr>
<tr>
<td></td>
<td>Out-of-sample prediction</td>
<td>21.84%</td>
<td>25.29%***</td>
<td>65.08%***</td>
<td>25.46%***</td>
<td>65.08%***</td>
</tr>
<tr>
<td></td>
<td>Surviving variables</td>
<td>268</td>
<td>143</td>
<td>45</td>
<td>144</td>
<td>45</td>
</tr>
<tr>
<td>EQIRs</td>
<td>In-sample prediction</td>
<td>93.17%</td>
<td>94.27%***</td>
<td>93.71%***</td>
<td>93.31%***</td>
<td>93.98%***</td>
</tr>
<tr>
<td></td>
<td>Out-of-sample prediction</td>
<td>48.98%</td>
<td>82.99%***</td>
<td>89.59%***</td>
<td>12.44%***</td>
<td>89.26%***</td>
</tr>
<tr>
<td></td>
<td>Surviving variables</td>
<td>268</td>
<td>167</td>
<td>95</td>
<td>181</td>
<td>87</td>
</tr>
</tbody>
</table>

Notes: The Table reports the Accuracy Ratios and the number of surviving variables for each model under study. “OP” stands for the ordered probit model. “OP_LASSO” refers to the ordered probit model with LASSO estimator. “OP_ELASTIC NET” stands for the ordered probit model with Elastic net estimator. “CR_LASSO” indicates the continuation ratio model with LASSO estimator. “CR_ELASTIC NET” is the continuation ratio model with Elastic net estimator. “AIC” is the AIC-type tuning parameter selector. “BIC” is the BIC-type tuning parameter selector. *** denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 1% significance level. ** denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 5% significance level. * denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 10% significance level.

Table 14: Accuracy Ratios and selected variables (random effects)

<table>
<thead>
<tr>
<th></th>
<th>OP_re</th>
<th>OP_re_LASSO</th>
<th>OP_re_ELASTIC NET</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AIC</td>
<td>BIC</td>
<td>AIC</td>
</tr>
<tr>
<td>CDSRs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-sample prediction</td>
<td>83.62%</td>
<td>84.15%***</td>
<td>83.00%***</td>
</tr>
<tr>
<td>Out-of-sample prediction</td>
<td>21.84%</td>
<td>25.29%***</td>
<td>65.08%***</td>
</tr>
<tr>
<td>Surviving variables</td>
<td>268</td>
<td>143</td>
<td>45</td>
</tr>
<tr>
<td>EQIRs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-sample prediction</td>
<td>93.17%</td>
<td>94.27%***</td>
<td>93.71%***</td>
</tr>
<tr>
<td>Out-of-sample prediction</td>
<td>48.98%</td>
<td>82.99%***</td>
<td>89.59%***</td>
</tr>
<tr>
<td>Surviving variables</td>
<td>268</td>
<td>167</td>
<td>95</td>
</tr>
</tbody>
</table>

Notes: The Table reports the Accuracy Ratios and the number of surviving variables for each model under study. “OP_re” stands for the ordered probit model with random effects. “OP_re_LASSO” is the ordered probit model with LASSO estimator and random effects. “OP_re_ELASTIC NET” stands for the ordered probit model with Elastic net estimator and random effects. “AIC” is the AIC-type tuning parameter selector. “BIC” is the BIC-type tuning parameter selector. *** denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 1% significance level. ** denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 5% significance level. * denotes that the Stuart-Maxwell null hypothesis of no difference of each category between two predictions is rejected at the 10% significance level.
## Table A1: Expected signs and variables definition

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Predicted relationship</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Firm-specific variables (16)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DETA</td>
<td>+</td>
<td>Logarithm of real total assets</td>
</tr>
<tr>
<td>Leverage (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AE</td>
<td>-</td>
<td>Total assets/Equity</td>
</tr>
<tr>
<td>LDA</td>
<td>-</td>
<td>Long-term debt/Total assets</td>
</tr>
<tr>
<td>SDA</td>
<td>-</td>
<td>Short-term debt/Total assets</td>
</tr>
<tr>
<td>TDA</td>
<td>-</td>
<td>Total debt/Total assets</td>
</tr>
<tr>
<td>TDEBITDA</td>
<td>-</td>
<td>Total debt/Earnings before interest, taxes, depreciation, amortization, and restructuring or rent costs</td>
</tr>
<tr>
<td>Coverage (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EBITINT</td>
<td>+</td>
<td>Earnings before interest and tax/Interest expenses</td>
</tr>
<tr>
<td>EBITDAINT</td>
<td>+</td>
<td>Total debt/earnings before interest, taxes, depreciation, amortization, and restructuring or rent costs/Interest expenses</td>
</tr>
<tr>
<td>Cash flow (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFOA</td>
<td>+</td>
<td>Cash flow from operating activities/Total assets</td>
</tr>
<tr>
<td>CASHEQA</td>
<td>+</td>
<td>Cash and equivalent/Total assets</td>
</tr>
<tr>
<td>Profitability (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OM</td>
<td>+</td>
<td>Operating income/Net sales</td>
</tr>
<tr>
<td>ROC</td>
<td>+</td>
<td>Net income less dividends/Total capital</td>
</tr>
<tr>
<td>ROE</td>
<td>+</td>
<td>Net income/Shareholders’ equity</td>
</tr>
<tr>
<td>ROA</td>
<td>+</td>
<td>Net income/Total assets</td>
</tr>
<tr>
<td>FFD</td>
<td>+</td>
<td>Funds from operations/Total debt</td>
</tr>
<tr>
<td>Liquidity (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIQ</td>
<td>+</td>
<td>Cash from operations/Liabilities</td>
</tr>
<tr>
<td>Market-driven Variables (6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXRET</td>
<td>+</td>
<td>Monthly stock return-the S&amp;P 500 index return</td>
</tr>
<tr>
<td>RSIZE</td>
<td>+</td>
<td>Firm equity value/Total market equity value</td>
</tr>
<tr>
<td>STA</td>
<td>-</td>
<td>The standard deviation of a company’s monthly stock returns</td>
</tr>
<tr>
<td>BETA</td>
<td>-</td>
<td>Systematic risk in the Capital Asset Pricing Model</td>
</tr>
<tr>
<td>PD1</td>
<td>-</td>
<td>1-year default probability</td>
</tr>
<tr>
<td>PD5</td>
<td>-</td>
<td>5-year default probability</td>
</tr>
<tr>
<td>Macroeconomic Variables (11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RLSP</td>
<td>~</td>
<td>Return on S&amp;P 500 index</td>
</tr>
<tr>
<td>CPFMM</td>
<td>~</td>
<td>3-month commercial paper rate</td>
</tr>
<tr>
<td>TB3</td>
<td>~</td>
<td>3-month Treasury bill rate minus federal funds rate</td>
</tr>
<tr>
<td>GS1</td>
<td>~</td>
<td>1-year constant maturity treasury rate</td>
</tr>
<tr>
<td>MB</td>
<td>~</td>
<td>Growth rate in the narrow money stock</td>
</tr>
<tr>
<td>INFL</td>
<td>~</td>
<td>Inflation rate</td>
</tr>
<tr>
<td>DLIP</td>
<td>~</td>
<td>Rate of change in industrial production</td>
</tr>
<tr>
<td>DLGDP</td>
<td>~</td>
<td>Real GDP growth</td>
</tr>
<tr>
<td>CFNA</td>
<td>~</td>
<td>Average Chicago Fed National Activity Index</td>
</tr>
<tr>
<td>UNRATE</td>
<td>~</td>
<td>Average unemployment rate</td>
</tr>
<tr>
<td>VIX</td>
<td>~</td>
<td>The Chicago Board Options Exchange volatility index</td>
</tr>
</tbody>
</table>

Notes: “+” indicates that the Market Implied Ratings would improve if the covariates rise. “-” indicates that the Market Implied Ratings would worsen if the covariates rise. “~” indicates uncertainty in the sign.
### On-line Appendix B

We cross tabulate predicted against observed CDSIRs outcomes in contingency Table B1 to B12 for the in-sample prediction.

#### Table B1: In-sample Prediction in Ordered Probit Model in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>Actual CDSIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>64</td>
<td>169</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>AA</td>
<td>34</td>
<td>879</td>
<td>71</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>995</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>45</td>
<td>1591</td>
<td>69</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1705</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>1</td>
<td>67</td>
<td>1804</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>1895</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>783</td>
<td>8</td>
<td>0</td>
<td>841</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>133</td>
<td>4</td>
<td>148</td>
<td>148</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>98</td>
<td>1094</td>
<td>1730</td>
<td>1934</td>
<td>817</td>
<td>146</td>
<td>6</td>
<td>5825</td>
</tr>
</tbody>
</table>

AR= 90.23%
CP= 69.29%

#### Table B2: In-sample Prediction in Ordered Probit Model with LASSO by AIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>Actual CDSIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>59</td>
<td>174</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>AA</td>
<td>25</td>
<td>887</td>
<td>70</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>995</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>45</td>
<td>1592</td>
<td>68</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1705</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>1</td>
<td>67</td>
<td>1805</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>1895</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>53</td>
<td>781</td>
<td>7</td>
<td>0</td>
<td>841</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>135</td>
<td>2</td>
<td>148</td>
<td>148</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>85</td>
<td>1107</td>
<td>1730</td>
<td>1938</td>
<td>814</td>
<td>148</td>
<td>3</td>
<td>5825</td>
</tr>
</tbody>
</table>

AR= 90.30%
CP= 66.89%

#### Table B3: In-sample Prediction in Ordered Probit Model with LASSO by BIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>Actual CDSIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>6</td>
<td>227</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>AA</td>
<td>6</td>
<td>905</td>
<td>56</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>995</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>41</td>
<td>1567</td>
<td>97</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1705</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>1</td>
<td>60</td>
<td>1813</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>1895</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>55</td>
<td>779</td>
<td>7</td>
<td>0</td>
<td>841</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>137</td>
<td>0</td>
<td>148</td>
<td>148</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>1174</td>
<td>1684</td>
<td>1993</td>
<td>812</td>
<td>151</td>
<td>0</td>
<td>5825</td>
</tr>
</tbody>
</table>

AR= 89.39%
CP= 61.04%
Table B4: In-sample Prediction in Ordered Probit Model with Elastic Net by AIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>Actual CDSRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>60</td>
<td>173</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>AA</td>
<td>27</td>
<td>886</td>
<td>70</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>995</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>45</td>
<td>1592</td>
<td>68</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1705</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>1</td>
<td>67</td>
<td>1805</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1895</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>53</td>
<td>781</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>841</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>135</td>
<td>2</td>
<td>148</td>
<td></td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>7</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>87</td>
<td>1105</td>
<td>1730</td>
<td>1938</td>
<td>1985</td>
<td>814</td>
<td>148</td>
<td>3</td>
<td>5825</td>
</tr>
</tbody>
</table>

AR= 90.30%
CP= 66.95%

Table B5: In-sample Prediction in Ordered Probit Model with Elastic Net by BIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>Actual CDSRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>6</td>
<td>227</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>AA</td>
<td>6</td>
<td>905</td>
<td>56</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>995</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>41</td>
<td>1567</td>
<td>97</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1705</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>1813</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1895</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>55</td>
<td>779</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>841</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>137</td>
<td>0</td>
<td>148</td>
<td></td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>1173</td>
<td>1684</td>
<td>1993</td>
<td>812</td>
<td>151</td>
<td>0</td>
<td>5825</td>
<td></td>
</tr>
</tbody>
</table>

AR= 89.39%
CP= 61.05%

Table B6: In-sample Prediction in Continuation Ratio Model with LASSO by AIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>Actual CDSRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>40</td>
<td>193</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>AA</td>
<td>18</td>
<td>893</td>
<td>61</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>995</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>42</td>
<td>1,569</td>
<td>94</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1705</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>1,813</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1895</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>55</td>
<td>779</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>841</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>137</td>
<td>0</td>
<td>148</td>
<td></td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>58</td>
<td>1128</td>
<td>1691</td>
<td>1985</td>
<td>812</td>
<td>151</td>
<td>0</td>
<td>5825</td>
<td></td>
</tr>
</tbody>
</table>

AR= 89.80%
CP= 63.29%
### Table B7: In-sample Prediction in Continuation Ratio Model with LASSO by BIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Actual CDSIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>19</td>
<td>214</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>AA</td>
<td>14</td>
<td>897</td>
<td>55</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>995</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>41</td>
<td>1567</td>
<td>97</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1705</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>1813</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>1895</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>55</td>
<td>779</td>
<td>7</td>
<td>0</td>
<td>841</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>137</td>
<td>0</td>
<td>148</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>1152</td>
<td>1683</td>
<td>1994</td>
<td>812</td>
<td>151</td>
<td>0</td>
<td>5825</td>
</tr>
</tbody>
</table>

AR= 89.48%
CP= 61.84%

### Table B8: In-sample Prediction in Continuation Ratio Model with Elastic Net by AIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Actual CDSIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>33</td>
<td>200</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>AA</td>
<td>21</td>
<td>890</td>
<td>61</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>995</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>41</td>
<td>1569</td>
<td>95</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1705</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>1813</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>1895</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>55</td>
<td>779</td>
<td>7</td>
<td>0</td>
<td>841</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>137</td>
<td>0</td>
<td>148</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>54</td>
<td>1131</td>
<td>1691</td>
<td>1986</td>
<td>812</td>
<td>151</td>
<td>0</td>
<td>5825</td>
</tr>
</tbody>
</table>

AR= 89.63%
CP= 62.74%

### Table B9: In-sample Prediction in Continuation Ratio Model with Elastic Net by BIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Actual CDSIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>19</td>
<td>214</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>AA</td>
<td>14</td>
<td>897</td>
<td>56</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>995</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>41</td>
<td>1567</td>
<td>97</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1705</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>1813</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>1895</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>55</td>
<td>779</td>
<td>7</td>
<td>0</td>
<td>841</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>137</td>
<td>0</td>
<td>148</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>1152</td>
<td>1684</td>
<td>1993</td>
<td>812</td>
<td>151</td>
<td>0</td>
<td>5825</td>
</tr>
</tbody>
</table>

AR= 89.48%
CP= 61.84%
Table B10: In-sample Prediction in Principal Component Analysis _16 PCs with Ordered Probit model in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>Actual CDSRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>37</td>
<td>156</td>
<td>40</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>AA</td>
<td>35</td>
<td>271</td>
<td>446</td>
<td>243</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>995</td>
</tr>
<tr>
<td>A</td>
<td>3</td>
<td>200</td>
<td>774</td>
<td>721</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1,705</td>
</tr>
<tr>
<td>BBB</td>
<td>5</td>
<td>33</td>
<td>536</td>
<td>1,255</td>
<td>64</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1,895</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>1</td>
<td>55</td>
<td>433</td>
<td>304</td>
<td>48</td>
<td>0</td>
<td>0</td>
<td>841</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>39</td>
<td>66</td>
<td>35</td>
<td>3</td>
<td>0</td>
<td>148</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>80</td>
<td>661</td>
<td>1,856</td>
<td>2,693</td>
<td>444</td>
<td>88</td>
<td>3</td>
<td>5825</td>
<td></td>
</tr>
</tbody>
</table>

AR= 45.94%

Commented [Office1]: Should I add CP ratio in following tables related PCA?

Table B11: In-sample Prediction in Principal Component Analysis _25 PCs with Ordered Probit model in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>Actual CDSRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>31</td>
<td>186</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>AA</td>
<td>42</td>
<td>440</td>
<td>420</td>
<td>93</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>995</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>184</td>
<td>873</td>
<td>626</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1,705</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>9</td>
<td>619</td>
<td>1,188</td>
<td>74</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1,895</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>356</td>
<td>417</td>
<td>57</td>
<td>2</td>
<td>0</td>
<td>841</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>39</td>
<td>81</td>
<td>43</td>
<td>5</td>
<td>0</td>
<td>148</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>74</td>
<td>819</td>
<td>1,938</td>
<td>2,282</td>
<td>597</td>
<td>108</td>
<td>7</td>
<td>5,825</td>
<td></td>
</tr>
</tbody>
</table>

AR= 51.36%

Table B12: In-sample Prediction in Principal Component Analysis _45 PCs with Ordered Probit model in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>Actual CDSRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>47</td>
<td>184</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>AA</td>
<td>70</td>
<td>483</td>
<td>397</td>
<td>45</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>995</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>109</td>
<td>1,027</td>
<td>550</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1,705</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>17</td>
<td>582</td>
<td>1,240</td>
<td>54</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1,895</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>245</td>
<td>514</td>
<td>68</td>
<td>4</td>
<td>0</td>
<td>841</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>37</td>
<td>77</td>
<td>32</td>
<td>2</td>
<td>0</td>
<td>148</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>793</td>
<td>2,019</td>
<td>2,117</td>
<td>666</td>
<td>105</td>
<td>6</td>
<td>5,825</td>
<td></td>
</tr>
</tbody>
</table>

AR= 57.39%
We cross tabulate predicted against observed CDSIRs outcomes in contingency Table B13 to B24 for the out-of-sample prediction.

### Table B13: Out-of-sample Prediction in Ordered Probit Model in CDSIRs

<table>
<thead>
<tr>
<th>Actual CDSIRs</th>
<th>Predicted CDSIRs</th>
<th>Below CCCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>9 14 4 1 23 22 11</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>24 38 51 8 43 48 22</td>
<td>234</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>1 1 35 50 65 8 42 45</td>
<td>246</td>
<td></td>
</tr>
<tr>
<td>BBB</td>
<td>0 0 0 36 82 58 40 139</td>
<td>355</td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td>0 0 0 0 10 37 14 76</td>
<td>137</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0 0 0 0 0 5 13 34</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Below CCC</td>
<td>0 0 0 0 0 3 20 23</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>34 87 141 166 174</td>
<td>347</td>
<td>1131</td>
</tr>
</tbody>
</table>

AR= 22.02%  
CP= 18.22%

### Table B14: Out-of-sample Prediction in Ordered Probit Model with LASSO by AIC-Typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Actual CDSIRs</th>
<th>Predicted CDSIRs</th>
<th>Below CCCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0 84 0 0 0 0 0 84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0 212 22 0 0 0 0 234</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0 12 224 10 0 0 0 246</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBB</td>
<td>0 0 12 340 3 0 0 355</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td>0 0 0 5 132 0 0 137</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0 0 0 0 4 48 0 52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below CCC</td>
<td>0 0 0 0 0 23 0 23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0 308 258 355 139 71 0 1131</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AR= 31.56%  
CP= 25.49%

### Table B15: Out-of-sample Prediction in Ordered Probit Model with LASSO by BIC-Typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Actual CDSIRs</th>
<th>Predicted CDSIRs</th>
<th>Below CCCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0 84 0 0 0 0 0 84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0 212 22 0 0 0 0 234</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0 12 224 10 0 0 0 246</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBB</td>
<td>0 0 12 340 3 0 0 355</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td>0 0 0 5 132 0 0 137</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0 0 0 0 4 48 0 52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below CCC</td>
<td>0 0 0 0 0 23 0 23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0 308 258 355 139 71 0 1131</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AR= 84.53%  
CP= 61.01%
Table B16: Out-of-sample Prediction in Ordered Probit Model with Elastic Net by AIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>1</td>
<td>20</td>
<td>15</td>
<td>22</td>
<td>20</td>
<td>6</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>62</td>
<td>64</td>
<td>49</td>
<td>39</td>
<td>20</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>1</td>
<td>94</td>
<td>71</td>
<td>45</td>
<td>33</td>
<td>2</td>
<td>246</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>107</td>
<td>109</td>
<td>75</td>
<td>63</td>
<td>355</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>57</td>
<td>20</td>
<td>60</td>
<td>137</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>17</td>
<td>34</td>
<td>52</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>19</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>83</td>
<td>174</td>
<td>249</td>
<td>271</td>
<td>175</td>
<td>178</td>
<td>1131</td>
</tr>
</tbody>
</table>

AR= 31.56%
CP= 25.49%

Table B17: Out-of-sample Prediction in Ordered Probit Model with Elastic Net by BIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>84</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>212</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>12</td>
<td>224</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>246</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>340</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>355</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>132</td>
<td>0</td>
<td>0</td>
<td>137</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>48</td>
<td>0</td>
<td>0</td>
<td>52</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>308</td>
<td>258</td>
<td>355</td>
<td>139</td>
<td>71</td>
<td>0</td>
<td>1131</td>
</tr>
</tbody>
</table>

AR= 84.53%
CP= 61.01%

Table B18: Out-of-sample Prediction in Continuation Ratio Model with LASSO by AIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>3</td>
<td>42</td>
<td>39</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>137</td>
<td>89</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>2</td>
<td>170</td>
<td>73</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>246</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>218</td>
<td>131</td>
<td>0</td>
<td>0</td>
<td>355</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>93</td>
<td>42</td>
<td>0</td>
<td>137</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>32</td>
<td>19</td>
<td>2</td>
<td>52</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>12</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>181</td>
<td>304</td>
<td>301</td>
<td>226</td>
<td>85</td>
<td>31</td>
<td>1131</td>
</tr>
</tbody>
</table>

AR=58.80%
CP=45.70%
### Table B19: Out-of-sample Prediction in Continuation Ratio Model with LASSO by BIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>Actual CDSIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>2</td>
<td>81</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>206</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>12</td>
<td>215</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>246</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>339</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>355</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>133</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>137</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>46</td>
<td>2</td>
<td>2</td>
<td>52</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>6</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>299</td>
<td>254</td>
<td>362</td>
<td>143</td>
<td>63</td>
<td>8</td>
<td>1131</td>
<td></td>
</tr>
</tbody>
</table>

AR= 83.73%
CP= 64.16%

### Table B20: Out-of-sample Prediction in Continuation Ratio Model with Elastic Net by AIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>Actual CDSIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>3</td>
<td>46</td>
<td>35</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>142</td>
<td>87</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>2</td>
<td>173</td>
<td>70</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>246</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>227</td>
<td>122</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>355</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>98</td>
<td>37</td>
<td>0</td>
<td>0</td>
<td>137</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>34</td>
<td>17</td>
<td>12</td>
<td>52</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>12</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>190</td>
<td>301</td>
<td>304</td>
<td>222</td>
<td>82</td>
<td>29</td>
<td>1131</td>
<td></td>
</tr>
</tbody>
</table>

AR= 60.92%
CP= 47.94%

### Table B21: Out-of-sample Prediction in Continuation Ratio Model with Elastic Net by BIC-typing Tuning Parameter Selector in CDSIRs

<table>
<thead>
<tr>
<th>Predicted CDSIRs</th>
<th>Actual CDSIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>2</td>
<td>81</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>209</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>12</td>
<td>218</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>246</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>341</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>355</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>132</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>137</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>47</td>
<td>1</td>
<td>52</td>
<td>4</td>
<td>23</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>19</td>
<td>4</td>
<td>23</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>302</td>
<td>254</td>
<td>362</td>
<td>140</td>
<td>66</td>
<td>5</td>
<td>1131</td>
<td></td>
</tr>
</tbody>
</table>

AR= 84.26%
CP= 63.42%
Table B22: Out-of-sample Prediction in Principal Component Analysis _16 PCs with Ordered Probit model in CDSIRs

<table>
<thead>
<tr>
<th>Actual CDSRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>15</td>
<td>47</td>
<td>21</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>AA</td>
<td>2</td>
<td>85</td>
<td>122</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>A</td>
<td>4</td>
<td>24</td>
<td>145</td>
<td>72</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>246</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>10</td>
<td>124</td>
<td>211</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>355</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>70</td>
<td>53</td>
<td>2</td>
<td>0</td>
<td>137</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>11</td>
<td>31</td>
<td>8</td>
<td>0</td>
<td>52</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>13</td>
<td>7</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>166</td>
<td>426</td>
<td>393</td>
<td>108</td>
<td>17</td>
<td>0</td>
<td>131</td>
</tr>
</tbody>
</table>

AR= 45.71%

Table B23: Out-of-sample Prediction in Principal Component Analysis _25 PCs with Ordered Probit model in CDSIRs

<table>
<thead>
<tr>
<th>Actual CDSRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>67</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>AA</td>
<td>89</td>
<td>130</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>A</td>
<td>45</td>
<td>98</td>
<td>81</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>246</td>
</tr>
<tr>
<td>BBB</td>
<td>46</td>
<td>114</td>
<td>126</td>
<td>64</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>355</td>
</tr>
<tr>
<td>BB</td>
<td>4</td>
<td>29</td>
<td>21</td>
<td>64</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td>137</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td>17</td>
<td>22</td>
<td>2</td>
<td>1</td>
<td>52</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>8</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>251</td>
<td>393</td>
<td>248</td>
<td>172</td>
<td>55</td>
<td>11</td>
<td>1</td>
<td>1131</td>
</tr>
</tbody>
</table>

AR= 32.01%

Table B24: Out-of-sample Prediction in Principal Component Analysis _45 PCs with Ordered Probit model in CDSIRs

<table>
<thead>
<tr>
<th>Actual CDSRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>6</td>
<td>42</td>
<td>29</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>77</td>
<td>108</td>
<td>49</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>234</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>12</td>
<td>99</td>
<td>131</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>246</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>2</td>
<td>44</td>
<td>237</td>
<td>70</td>
<td>2</td>
<td>0</td>
<td>355</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>29</td>
<td>87</td>
<td>18</td>
<td>3</td>
<td>137</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>31</td>
<td>11</td>
<td>4</td>
<td>52</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>6</td>
<td>133</td>
<td>280</td>
<td>459</td>
<td>200</td>
<td>39</td>
<td>14</td>
<td>1131</td>
</tr>
</tbody>
</table>

AR= 46.33%
We cross tabulate predicted against observed EQIRs outcomes in contingency Table B25 to B36 for the in-sample prediction.

### Table B25: In-sample Prediction in Ordered Probit Model in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>3</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>409</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>441</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>19</td>
<td>2492</td>
<td>115</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2626</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>2915</td>
<td>56</td>
<td>0</td>
<td>0</td>
<td>3041</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>46</td>
<td>1273</td>
<td>14</td>
<td>0</td>
<td>1333</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>173</td>
<td>0</td>
<td>187</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>440</td>
<td>2594</td>
<td>3076</td>
<td>1343</td>
<td>187</td>
<td>0</td>
<td>7643</td>
</tr>
</tbody>
</table>

AR: 95.05%
CP: 78.30%

### Table B26: In-sample Prediction in Ordered Probit Model with LASSO by AIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>2</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>410</td>
<td>31</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>441</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>16</td>
<td>2495</td>
<td>115</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2626</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>72</td>
<td>2910</td>
<td>59</td>
<td>0</td>
<td>0</td>
<td>3041</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>1269</td>
<td>14</td>
<td>0</td>
<td>1333</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>175</td>
<td>0</td>
<td>187</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>439</td>
<td>2598</td>
<td>3075</td>
<td>1340</td>
<td>189</td>
<td>0</td>
<td>7643</td>
</tr>
</tbody>
</table>

AR: 95.00%
CP: 77.16%

### Table B27: In-sample Prediction in Ordered Probit Model with LASSO by BIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>412</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>441</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>18</td>
<td>2492</td>
<td>116</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2626</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>84</td>
<td>2895</td>
<td>62</td>
<td>0</td>
<td>0</td>
<td>3041</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>55</td>
<td>1262</td>
<td>16</td>
<td>0</td>
<td>1333</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>176</td>
<td>0</td>
<td>0</td>
<td>187</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>445</td>
<td>2605</td>
<td>3066</td>
<td>1335</td>
<td>192</td>
<td>0</td>
<td>7643</td>
</tr>
</tbody>
</table>

AR: 94.69%
CP: 74.46%
### Table B28: In-sample Prediction in Ordered Probit Model with Elastic Net by AIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Predicted EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>2</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>410</td>
<td>31</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>441</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>16</td>
<td>2498</td>
<td>112</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2626</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>71</td>
<td>2911</td>
<td>59</td>
<td>0</td>
<td>0</td>
<td>3041</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>49</td>
<td>1269</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>1333</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>175</td>
<td>0</td>
<td>0</td>
<td>187</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>439</td>
<td>2600</td>
<td>3072</td>
<td>1340</td>
<td>190</td>
<td>0</td>
<td>7643</td>
</tr>
</tbody>
</table>

AR=95.05%
CP=77.19%

### Table B29: In-sample Prediction in Ordered Probit Model with Elastic Net by BIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Predicted EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>412</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>441</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>20</td>
<td>2490</td>
<td>116</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2626</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>92</td>
<td>2888</td>
<td>61</td>
<td>0</td>
<td>0</td>
<td>3041</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>58</td>
<td>1259</td>
<td>16</td>
<td>0</td>
<td>1333</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>175</td>
<td>0</td>
<td>0</td>
<td>187</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>447</td>
<td>2611</td>
<td>3062</td>
<td>1332</td>
<td>191</td>
<td>0</td>
<td>7643</td>
</tr>
</tbody>
</table>

AR=94.52%
CP=74.25%

### Table B30: In-sample Prediction in Continuation Ratio Model with LASSO by AIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Predicted EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>3</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>409</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>441</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>15</td>
<td>2501</td>
<td>110</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2626</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>72</td>
<td>2914</td>
<td>55</td>
<td>0</td>
<td>0</td>
<td>3041</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>47</td>
<td>1273</td>
<td>13</td>
<td>0</td>
<td>1333</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>175</td>
<td>0</td>
<td>0</td>
<td>187</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>436</td>
<td>2605</td>
<td>3071</td>
<td>1340</td>
<td>188</td>
<td>0</td>
<td>7643</td>
</tr>
</tbody>
</table>

AR=95.19%
CP=78.58%
Table B31: In-sample Prediction in Continuation Ratio Model with LASSO by BIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Predicted EQIRs</th>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below</th>
<th>CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>412</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>441</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>15</td>
<td>2495</td>
<td>116</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2626</td>
<td></td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>82</td>
<td>2897</td>
<td>62</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3041</td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>55</td>
<td>1263</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1333</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>176</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>187</td>
<td></td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>442</td>
<td>2606</td>
<td>3068</td>
<td>1336</td>
<td>191</td>
<td>0</td>
<td>0</td>
<td>7643</td>
<td></td>
</tr>
</tbody>
</table>

AR=94.77%
CP=74.51%

Table B32: In-sample Prediction in Continuation Ratio Model with Elastic Net by AIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Predicted EQIRs</th>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below</th>
<th>CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>3</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>409</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>441</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>16</td>
<td>2497</td>
<td>113</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2626</td>
<td></td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>2918</td>
<td>53</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3041</td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>49</td>
<td>1272</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>1333</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>174</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>187</td>
<td></td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>437</td>
<td>2599</td>
<td>3080</td>
<td>1338</td>
<td>186</td>
<td>0</td>
<td>0</td>
<td>7643</td>
<td></td>
</tr>
</tbody>
</table>

AR=95.16%
CP=78.45%

Table B33: In-sample Prediction in Continuation Ratio Model with Elastic Net by BIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Predicted EQIRs</th>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below</th>
<th>CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>412</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>441</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>16</td>
<td>2494</td>
<td>116</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2626</td>
<td></td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>83</td>
<td>2896</td>
<td>62</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3041</td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>56</td>
<td>1263</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>1333</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>175</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>187</td>
<td></td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>443</td>
<td>2606</td>
<td>3068</td>
<td>1337</td>
<td>189</td>
<td>0</td>
<td>0</td>
<td>7643</td>
<td></td>
</tr>
</tbody>
</table>

AR=94.73%
CP=74.39%
### Table B34: In-sample Prediction in Principal Component Analysis - 15 PCs with Ordered Probit model in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>14</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>AA</td>
<td>3</td>
<td>152</td>
<td>278</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>441</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>148</td>
<td>1651</td>
<td>817</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>2626</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>4</td>
<td>559</td>
<td>2318</td>
<td>160</td>
<td>0</td>
<td>0</td>
<td>3041</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>612</td>
<td>674</td>
<td>47</td>
<td>0</td>
<td>1333</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>113</td>
<td>67</td>
<td>0</td>
<td>187</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>318</td>
<td>2489</td>
<td>3762</td>
<td>957</td>
<td>114</td>
<td>0</td>
<td>7643</td>
</tr>
</tbody>
</table>

AR = 63.61%

### Table B35: In-sample Prediction in Principal Component Analysis - 23 PCs with Ordered Probit model in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>174</td>
<td>258</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>441</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>102</td>
<td>1929</td>
<td>589</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>2626</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>1</td>
<td>433</td>
<td>2440</td>
<td>164</td>
<td>3</td>
<td>0</td>
<td>3041</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>408</td>
<td>864</td>
<td>61</td>
<td>0</td>
<td>1333</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>86</td>
<td>96</td>
<td>0</td>
<td>187</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>292</td>
<td>2620</td>
<td>3451</td>
<td>1119</td>
<td>160</td>
<td>0</td>
<td>7643</td>
</tr>
</tbody>
</table>

AR = 72.00%

### Table B36: In-sample Prediction in Principal Component Analysis - 41 PCs with Ordered Probit model in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>286</td>
<td>154</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>441</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>74</td>
<td>2052</td>
<td>500</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2626</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>393</td>
<td>2465</td>
<td>183</td>
<td>0</td>
<td>0</td>
<td>3041</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>272</td>
<td>990</td>
<td>71</td>
<td>0</td>
<td>1333</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>91</td>
<td>87</td>
<td>0</td>
<td>187</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>375</td>
<td>2599</td>
<td>3247</td>
<td>1264</td>
<td>158</td>
<td>0</td>
<td>7643</td>
</tr>
</tbody>
</table>

AR = 76.93%
We cross tabulate predicted against observed EQIRs outcomes in contingency Table B37 to B48 for the out-of-sample prediction.

### Table B37: Out-of-sample Prediction in Ordered Probit Model in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>Predicted EQIRs</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0 1 0 0 0 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>7 51 45 11 0 0 0 114</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>28 13 248 172 38 0 0 499</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBB</td>
<td>11 26 17 185 123 34 0 396</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td>0 9 14 11 78 40 0 152</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0 0 0 1 2 17 0 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below CCC</td>
<td>0 0 0 0 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>46 100 324 380 241 91 0 1182</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AR= 48.98%
CP= 35.49%

### Table B38: Out-of-sample Prediction in Ordered Probit Model with LASSO by AIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>Predicted EQIRs</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0 1 0 0 0 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0 93 21 0 0 0 0 114</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0 7 434 58 0 0 0 499</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBB</td>
<td>0 0 21 334 41 0 0 396</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td>0 0 0 17 121 14 0 152</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0 0 0 0 0 20 0 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below CCC</td>
<td>0 0 0 0 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0 101 476 409 162 34 0 1182</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AR=84.77%
CP=66.50%

### Table B39: Out-of-sample Prediction in Ordered Probit Model with LASSO by BIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>Predicted EQIRs</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0 1 0 0 0 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0 109 5 0 0 0 0 114</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0 8 475 16 0 0 0 499</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBB</td>
<td>0 0 39 349 8 0 0 396</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td>0 0 0 26 124 2 0 152</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0 0 0 0 2 18 0 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below CCC</td>
<td>0 0 0 0 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0 118 519 391 134 20 0 1182</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AR=90.95%
CP=70.10%
### Table B40: Out-of-sample Prediction in Ordered Probit Model with Elastic Net by AIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>85</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>114</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>5</td>
<td>411</td>
<td>82</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>499</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>319</td>
<td>59</td>
<td>0</td>
<td>0</td>
<td>396</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>121</td>
<td>15</td>
<td>0</td>
<td>152</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>91</td>
<td>458</td>
<td>417</td>
<td>181</td>
<td>35</td>
<td>0</td>
<td>1182</td>
</tr>
</tbody>
</table>

AR=90.88%
CP=63.42%

### Table B41: Out-of-sample Prediction in Ordered Probit Model with Elastic Net by BIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>109</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>114</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>9</td>
<td>474</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>499</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>39</td>
<td>349</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>396</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>27</td>
<td>123</td>
<td>2</td>
<td>0</td>
<td>152</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>18</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>119</td>
<td>518</td>
<td>392</td>
<td>133</td>
<td>20</td>
<td>0</td>
<td>1182</td>
</tr>
</tbody>
</table>

AR=90.78%
CP=69.93%

### Table B42: Out-of-sample Prediction in Continuation Ratio Model with LASSO by AIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>93</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>114</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>6</td>
<td>436</td>
<td>57</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>499</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>335</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>396</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>126</td>
<td>11</td>
<td>0</td>
<td>152</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>100</td>
<td>478</td>
<td>407</td>
<td>166</td>
<td>31</td>
<td>0</td>
<td>1182</td>
</tr>
</tbody>
</table>

AR=85.45%
CP=67.29%
### Table B43: Out-of-sample Prediction in Continuation Ratio Model with LASSO by BIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Predicted EQIRs</th>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>109</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>114</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>8</td>
<td>475</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>499</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>39</td>
<td>349</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>396</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>125</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>152</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>118</td>
<td>519</td>
<td>390</td>
<td>135</td>
<td>20</td>
<td>0</td>
<td>1182</td>
<td></td>
</tr>
</tbody>
</table>

AR=91.03%
CP=70.23%

### Table B44: Out-of-sample Prediction in Continuation Ratio Model with Elastic Net by AIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Predicted EQIRs</th>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>86</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>114</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>7</td>
<td>410</td>
<td>82</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>499</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>321</td>
<td>57</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>396</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>119</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>152</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>94</td>
<td>456</td>
<td>421</td>
<td>177</td>
<td>34</td>
<td>0</td>
<td>1182</td>
<td></td>
</tr>
</tbody>
</table>

AR=80.80%
CP=62.39%

### Table B45: Out-of-sample Prediction in Continuation Ratio Model with Elastic Net by BIC-typing Tuning Parameter Selector in EQIRs

<table>
<thead>
<tr>
<th>Predicted EQIRs</th>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>109</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>114</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>8</td>
<td>475</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>499</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>39</td>
<td>349</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>396</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td>124</td>
<td>2</td>
<td>0</td>
<td>20</td>
<td>152</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>118</td>
<td>519</td>
<td>391</td>
<td>134</td>
<td>20</td>
<td>0</td>
<td>1182</td>
<td></td>
</tr>
</tbody>
</table>

AR=90.95%
CP=70.10%
### Table B46: Out-of-sample Prediction in Principal Component Analysis with Ordered Probit model in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>51</td>
<td>61</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>114</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>18</td>
<td>357</td>
<td>119</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>499</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>84</td>
<td>269</td>
<td>43</td>
<td>0</td>
<td>0</td>
<td>396</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>80</td>
<td>65</td>
<td>0</td>
<td>0</td>
<td>152</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>12</td>
<td>7</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>70</td>
<td>509</td>
<td>471</td>
<td>125</td>
<td>7</td>
<td>0</td>
<td>1182</td>
</tr>
</tbody>
</table>

AR= 63.37%

### Table B47: Out-of-sample Prediction in Principal Component Analysis with Ordered Probit model in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>4</td>
<td>34</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>76</td>
<td>355</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>114</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>2</td>
<td>120</td>
<td>33</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>499</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>32</td>
<td>83</td>
<td>36</td>
<td>1</td>
<td>0</td>
<td>152</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>12</td>
<td>7</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>6</td>
<td>213</td>
<td>627</td>
<td>279</td>
<td>49</td>
<td>8</td>
<td>1182</td>
</tr>
</tbody>
</table>

AR= 54.74%

### Table B48: Out-of-sample Prediction in Principal Component Analysis with Ordered Probit model in EQIRs

<table>
<thead>
<tr>
<th>Actual EQIRs</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>Below CCC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0</td>
<td>10</td>
<td>80</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>114</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>2</td>
<td>150</td>
<td>323</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>499</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>14</td>
<td>210</td>
<td>171</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>396</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>92</td>
<td>35</td>
<td>0</td>
<td>0</td>
<td>152</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Below CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>12</td>
<td>245</td>
<td>582</td>
<td>291</td>
<td>50</td>
<td>2</td>
<td>1182</td>
</tr>
</tbody>
</table>

AR= 51.69%
On-line Appendix C

In this section, we report the procedure for selecting tuning parameters.

In our study, we select the tuning parameter with the aid of the AIC (Akaike, 1974) and the BIC (Schwartz, 1978). These are presented below:

\[ AIC = -2 \times \log(\text{maximum likelihood}) + 2 \times (\text{the number of used parameters within the model}) \]

and,

\[ BIC = -2 \times \log(\text{maximum likelihood}) + 2 \times (\text{the number of used parameters within the model}) \times \log(\text{the number of observations}) \]

The number of used parameters within the model is the number of non-zero estimated coefficients of used parameters (Efron et al., 2004 and Zou et al., 2007). We follow the parametrization procedure of Wurm et al., 2017.

The exact algorithm for the LASSO tuning parameter based on the AIC and the BIC criteria is presented below:

1. Estimate the intercept-only model by maximum likelihood (see, equation (4) for the OP with LASSO or equation (9) for the CR with LASSO).
2. Select a sequence of tuning parameters \( \lambda_i \) \((i = 1 \text{ to } n)\).
   For each tuning parameter \( \lambda_i \), there exists a set of selected independent variables and corresponding estimated coefficients. In other words, different potential models including different selected predictors controlled by the sequence of tuning parameters \( \lambda_1, \lambda_2, \ldots, \lambda_n \) are constructed.
3. Calculate the AIC or BIC for each potential model.
4. Choose the models with minimum value of AIC or BIC from the aforementioned series. These two models (one based on AIC-type tuning parameter selector and one based on BIC-type tuning parameter selector) are applied in our study.

Similarly, for the Elastic net estimator we follow Wurm et al., 2017:

1. Estimate the intercept-only model by maximum likelihood (see, equation (6) for the OP with Elastic net or equation (11) for the CR with Elastic net).
2. Select a sequence of tuning parameters \( \alpha_1, \alpha_2, \ldots, \alpha_{10} \).
3. For each \( \alpha_i \) \((i = 1 \text{ to } 10)\), a sequence of tuning parameters \( \lambda_i \) \((i = 1 \text{ to } n)\) is selected. For each tuning parameter \( \lambda_i \), there is a set of selected independent variables and the corresponding estimated coefficients. For each one of these models, we generate an AIC and BIC value.
4. The min AIC and BIC values are assigned to the corresponding \( \alpha_i \).
5. The \( \alpha_i \) \((i = 1 \text{ to } 10)\) with the minimum AIC and the one with the minimum BIC value are selected. These two models are applied in our study.