A Predictive Bayesian Model Averaging Approach on Firm Default Probabilities

Laura Vana[1], Bettina Grün[2], Paul Hofmarcher[2]

Second Bayesian Young Statisticians Meeting (BAYSM 2014)
Vienna, September 18–19, 2014

[1] WU Wirtschaftsuniversität Wien, Department of Finance, Accounting and Statistics, Vienna, Austria
laura.vana@wu.ac.at

[2] Johannes Kepler University Linz, Department of Applied Statistics, Linz, Austria
<bettina.gruen, paul.hofmarcher>@jku.at

Abstract

A range of different financial ratios is in general used to predict a company's ability to meet its financial obligations. However, due to lack of theory no preferred set of ratios is indicated to build a simple, but well performing predictive model. Using a predictive model averaging approach we aim at identifying such simple models where not only the total number of regressors is restricted, but also their assignment to different categories of the financial ratios. We also compare the performance of these models to the famous Altman's Z-score model which contains five financial ratios from four different categories.

Keywords: credit risk; financial ratios; Altman Z-score; pseudo-marginal likelihood; Bayesian bootstrap.

1 Introduction

The most common approach in credit risk modeling is to link financial ratios to the company's ability to repay its obligations (creditworthiness). Financial ratios can be classified into different categories (i.e., liquidity, profitability, etc.). Each category measures a different financial aspect of the firm. Altman's prominent Z-score model [1] uses five financial ratios from four categories to measure the financial health of a company, i.e., to predict bankruptcy. The five ratios are selected out of 22 potential variables without a theoretical foundation explaining this choice. Over the last decades, various other ratios have been proposed as relevant for credit risk and thus could potentially be used in such a predictive model.

We address this problem of model choice and selection of suitable ratios by collecting an extensive list of ratios and determining simple and interpretable models with good predictive performance which use a subset of these ratios. As financial ratios belonging to the same category are highly correlated, we impose restrictions regarding the selection of several ratios from the same category. In this context Bayesian model averaging (BMA) [4] provides the framework for estimating parsimonious models when the set of potential regressors is large. To avoid in-sample overfitting, we combine BMA with predictive measures of fit.

2 Methods

We use BMA in the linear regression setting to model the relationship between the variables of interest. We choose an improper prior for the error variance $p(\sigma) \propto \sigma^{-1}$ and Zellner's g-prior on the coefficients with $g = N$, the number of observations. We compare the models based on their predictive ability. Thus, the weights assigned to the competing models are derived from the pseudo-marginal likelihood (PML) rather than the standard marginal likelihood. Following [3], the PML of observation i is univariate t:

$$Y_i | x_i, y, X \sim t(N, x_i^\top m_*, a_0(1 + x_i^\top V_* x_i)/N)$$

with

$$\begin{aligned} a_0 &= y^\top y - \frac{g}{g+1} y^\top X (X^\top X)^{-1} X^\top y, \\ m_* &= \frac{g}{g+1} (X^\top X)^{-1} X^\top y, \\ V_* &= \left(X^\top X + \frac{1}{g} X^\top X \right)^{-1}, \end{aligned}$$

where X is a design matrix and y is the dependent variable, both after standardization of variables. We identify the 1000 models with the highest average PML. When computing the PML all observations are used for determining the estimates to reduce the computational expenses. The predictive performance of the models is evaluated based on the Bayesian bootstrap procedure [5] using 10000 Bayesian bootstrap samples. In one bootstrap iteration, $q = [q_i]_{i \in 1:N}$ is drawn from a Dirichlet distribution with parameter $\alpha = (\alpha_1, \dots, \alpha_N) = (1, \dots, 1)$. By weighting the log PML with the vector q we get the bootstrap replicate of the sample statistic. The posterior model probability for a model M_γ is calculated as the proportion of bootstrap samples for which M_γ has the highest log PML. The final estimates are then constructed as a weighted average of the parameter estimates from each model.

3 Data and results

We analyse 1446 US corporations between 2009–2013, with a total of 5342 observations. The proxy for firms' creditworthiness are issuer credit ratings from Standard&Poor's. We map the ordinal ratings to probabilities of default (PDs) by using historical default rates. The dependent variable is the probit of the PDs. The matrix of regressors contains 74 ratios computed from financial statement data, mainly following [1, 2]. We add two ratio categories to the ones in [1] and thus have: interest coverage (5 ratios), liquidity (11 ratios), capital structure (21 ratios), profitability (17 ratios), cash flow (6 ratios) and efficiency (14 ratios). To achieve models of similar complexity to Altman's Z-score, we restrict the model set to models which include at most one (or two) ratio(s) from each category. Table 1 contains posterior results for the variables included in Altman's Z-score model or which had a posterior inclusion probability (PIP) greater than 0.1 in the BMA approach. We observe that four Z-score ratios have PIPs greater than 0.15. This would imply that Altman's model has still its validity today. However, we also find other candidates with higher PIPs that seem to offer a better predictive performance. One such candidate is *debt/earnings before interest, tax, depreciation and amortization.* This ratio is reported to be highly used by financial specialists in practice.

Table 1: Statistics for the potential set of relevant ratios

	PIP	Posterior Mean	Posterior SD	Category
interest expense/assets	0.998	0.458	0.047	interest coverage
quick assets/liabilities	0.101	0.010	0.034	liquidity
current assets/current liabilities	0.421	0.068	0.085	liquidity
cash/liabilities	0.110	0.011	0.033	liquidity
working capital/assets[1]	0.166	0.017	0.041	liquidity
debt/earnings before interest, tax, depreciation and amortization	0.986	0.222	0.041	capital structure
equity/liabilities[1]	0.007	0.000	0.010	capital structure
retained earnings/assets[1]	0.956	−0.189	0.046	profitability
earnings before interest and tax/assets[1]	0.169	−0.034	0.081	profitability
earnings before interest and tax/sales	0.577	−0.123	0.109	profitability
cash flow/debt	0.173	0.014	0.034	cash flow
cash flow/sales	0.434	−0.065	0.082	cash flow
cash flow/assets	0.112	−0.007	0.021	cash flow
sales/assets[1]	0.024	0.000	0.019	efficiency
cash/expenditures for operations	0.265	0.038	0.065	efficiency
quick assets/expenditures for operations	0.550	0.089	0.085	efficiency

[1]Ratios in Altman's Z-score.

References

[1] Altman, E. I. (1968). "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy." *The Journal of Finance*, **23**(4), 589–609.

[2] Beaver, W. H. (1966). "Financial ratios as predictors of failure." *Journal of Accounting Research*, **4**, 71–111.

[3] Eklund, J. and Karlsson, S. (2007). "Forecast combination and model averaging using predictive measures." *Econometric Reviews*, **26**(2–4), 329–363.

[4] Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). "Bayesian model averaging: A tutorial." *Statistical Science*, **14**(4), 382–401.

[5] Jackson, C. H., Sharples, L. D., and Thompson, S. G. (2010). "Structural and parameter uncertainty in Bayesian cost-effectiveness models." *Journal of the Royal Statistical Society C*, **59**(2), 233–253.