Modelling Loss Given Default with ESG Information

Junfeng Zhang, Galina Andreeva, Yizhe Dong

Credit Research Centre, University of Edinburgh
J.Zhang-161@sms.ed.ac.uk

Credit Scoring and Credit Control Conference XVIII
Augest 31, 2023

CRC Credit Research Centre

Presentation Overview

Modelling LGD with ESG Information

Motivation

- Existing studies shows that show that a better ESG performance can effectively mitigate a firm's credit risk
- How does specific credit risk components such as probability of default (PD) and loss given default (LGD) interact with ESG is still unknown.
- Breaking the overall credit risk measure down and analyzing how ESG interact with these specific credit risk components are helpful with better understand how ESG affect credit risk, thus to better manage credit risk under the Basel Accord.

References

Key Literature

Jankowitsch, R., Nagler, F., Subrahmanyam, M. G. (2014)
The determinants of recovery rates in the US corporate bond market
Journal of Financial Economics 114(1), 155-177.

Yao, X., Crook, J., Andreeva, G. (2015)
Support vector regression for loss given default modelling
European Journal of Operational Research 240(2), 528-538.

Yao, X., Crook, J., Andreeva, G. (2017)
Enhancing two-stage modelling methodology for loss given default with support vector machines
European Journal of Operational Research 263(2), 679-689.

Henisz, W. J., McGlinch, J. (2019)
ESG, material credit events, and credit risk.
Journal of Applied Corporate Finance 31(2), 105-117.

Kellner, R., Nagl, M., Rösch, D. (2022)
Opening the black box–Quantile neural networks for loss given default prediction
Journal of Banking and Finance 134, 106334.

Contribution

- We prove that ESG information is useful in LGD modelling from a predictive paradigm.
- We expand this finding by exploring the impact of ESG variables on LGD in different segmentation models.
- We find the relationship between LGD and ESG variables follows a temporal structure with the rolling window regression, and ESG information are more effective for estimating LGD in adverse macroeconomic environments (e.g., financial crisis).

Data

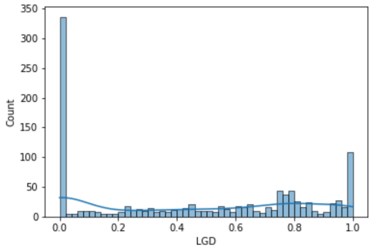
Data Sources

Type of Variables	Source	Access Platform
Macroeconomic Information	U.S. Bureau of Economic Analysis	U.S. St. Louis Federal Reserve
Financial Ratios	Compustat	Wharton Research Database Services
Debt-Level Characteristics	Moody's Default and Recovery Database	Moody
CSR/ESG Information	MSCI:KLD	Wharton Research Database Services
Loss Given Default Measurement	Moody's Default and Recovery Database	Moody

Table: Sources of Variables

Data
Data Highlights

- 1,104 Samples from 211 firms between 1992 and 2019
- 4 independent variable groups:
  - Debt-level characteristics: Seniority, Principal amount, etc.
  - Financial ratios: LTR, Intangible ratio, Profitability, etc.
  - Macroeconomic variables: UNRATE, DFF, etc.
  - ESG information
- Typical Bi-modal distribution of LGD



Data

ESG Variable Details

Perspective	**Involved Variables**
Environmental Concerns	ENV_con_num
Environmental Strengths	ENV_str_num
Social Concerns	COM_con_num, DIV_con_num, PRO_con_num, HUM_con_num, EMP_con_num, Other_con
Social Strengths	COM_str_num, DIV_str_num, PRO_str_num, EMP_str_num
Company Governance Concerns	CGOV_con_num
Company Governance Strengths	CGOV_str_num

Table: ESG Variables

Methodology
Truncated GBDT - 1

For dataset $\{(\mathbf{x}_i, y_i)\}_{i=1}^{n}$, and a differentiable loss function $L(y_i, F(\mathbf{x}))$ (squared error for this regression task), the gradient boosting tree regressor initiate its base learner with a constant value, i.e.,

$$F_0(\mathbf{x}) = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{n} L(y_i, \gamma) \tag{1}$$

Then, for $m = 1$ to M, the model compute

$$r_{im} = -\left[\frac{\partial L(y_i, F(\mathbf{x}_i))}{\partial F(\mathbf{x}_i)}\right]_{F(\mathbf{x})=F_{m-1}(\mathbf{x})} \tag{2}$$

for $i = 1, ..., n$ and fit a regression tree to the r_{im} values and create terminal regions R_{jm}.
For $j = 1, ..., J_m$, the model computes

$$\gamma_{jm} = \underset{\gamma}{\operatorname{argmin}} \sum_{\mathbf{x}_i \in R_{ij}} L(y_i, F_{m-1}(\mathbf{x}_i) + \gamma) \tag{3}$$

Methodology

Truncated GBDT - 2

At the end of each iteration for $m = 1$ to M, update

$$F_M(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \beta_m \sum_{j=1}^{J_m} \gamma_{jm} \, I(\mathbf{x} \in R_{jm}) \tag{4}$$

where β is the weight and the final predictor of gradient boosting tree regressor is $F(\mathbf{x})$.

On the top of this gradient boosting tree regressor, we use truncated setting to better estimate LGD. The truncation setting makes the prediction of LGD bounded in $[0, 1]$, i.e.,

$$\hat{y}_{\text{truncated}} = \min(1, \max(0, \hat{y})) \tag{5}$$

- The dataset is split 70:30 into training set and test set with out-of-time approach.
- Evaluation metrics: MSE and MAE

Methodology

AutoEncoder for Feature Extraction - 1

An AutoEncoder is a neural network model that seeks to learn a compressed representation of an input. Because a represented single score from irrelevant but proven key factors of LGD is needed to reduce the interference on the focused ESG variables.

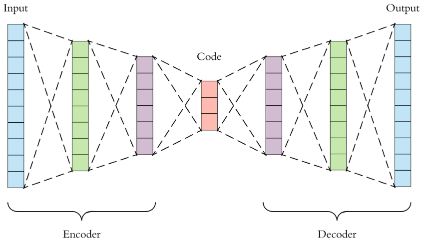


Figure: The Structure of AutoEncoder

Methodology

AutoEncoder for Feature Extraction - 2

Before feeding the data into the autoEncoder, the data must be scaled between 0 and 1 using MinMaxScaler, because we will use the Sigmoid activation function in the output layer, which outputs a value between 0 and 1.

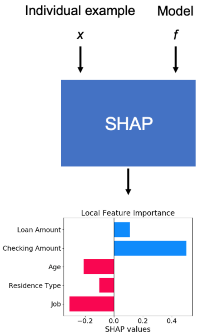
	Architecture
Encoder	Dense Layer, 32 neurons, Activation function = ReLu Dense Layer, 16 neurons, Activation function = ReLu Dense Layer, 8 neurons, Activation function = ReLu Dense Layer, 1 neurons, Activation function = ReLu
Decoder	Dense Layer, 8 neurons, Activation function = ReLu Dense Layer, 16 neurons, Activation function = ReLu Dense Layer, 32 neurons, Activation function = ReLu Dense Layer, # of features, Activation function = Sigmoid

Table: The Architechture of the AutoEncoder

Methodology

SHAP (SHapley Additive exPlanations) and Model Interpretability

- The goal of SHAP is to explain the prediction of an instance x by computing the contribution of each feature to the prediction.
- The SHAP explanation method computes Shapley values from coalitional game theory.
- SHAP is a model-agnostic method. SHAP TreeExplainer is especially compatible with tree-based learners with faster processing speed.



Result

R1. Main Model

Out-of-time Split	All	No ESG	No Fin	No Macro	ESG	Fin	Macro
MSE	0.1046***	0.1173***	0.0996***	0.1119***	0.1085***	0.1323***	0.1142***
MAE	0.2665***	0.2828***	0.2614***	0.2739***	0.2716***	0.3062***	0.2804***

Out-of-sample Split	All	No ESG	No Fin	No Macro	ESG	Fin	Macro
MSE	0.0449***	0.0458***	0.0492	0.0495	0.0693***	0.0498	0.0517***
MAE	0.1369***	0.1387***	0.1489***	0.1500***	0.1896***	0.1520***	0.1547***

Table: Modelling Performance with Different Variable Groups (***indicating all pairs with this model are significantly different at 99% level)

Result

R1. Main Model

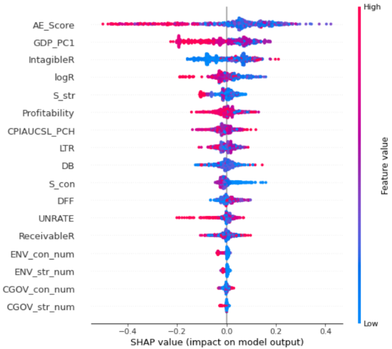


Figure: Summary Plot of SHAP Values

Result

R2. Segmentation Model 1 - Seniority

TYPE	Mean LGD	Count
Junior Subordinated Bonds	0.9746	7
Revolver	0.1201	169
Senior Secured Bonds	0.4544	217
Senior Subordinated Bonds	0.7934	43
Senior Unsecured Bonds	0.5763	477
Subordinated Bonds	0.7374	23
Term Loans	0.2846	168

Table: Mean LGD for Each Seniority Subgroup

	Secured	Secured w/o ESG	Unsecured	Unsecured w/o ESG
MSE	0.1983	0.1920	0.1076	0.1134
MAE	0.3596	0.3541	0.2685	0.2745

Table: Modelling Performance in Seniority Segmentation

Result

R2. Segmentation Model 1 - Seniority

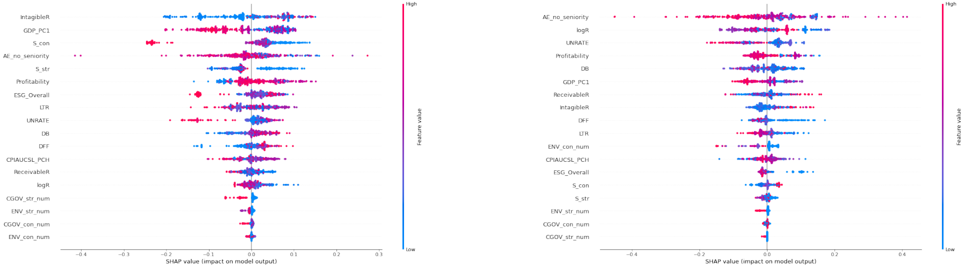


Figure: Summary Plots of SHAP, Unsecured (Left) vs Secured (Right)

Result

R3. Segmentation Model 2 - Industry

Segmentation	Industry	Count
Green	Telecommunications Media Retail Consumer Products Technology Services Distribution Packaging Gaming: Casinos Healthcare Restaurants Retail: Specialty Natural Products	436
Brown	Energy Transportation Automotive Manufacturing Chemicals Metal & Mining Construction Pharmaceuticals Aircraft & Aerospace	668

Table: Count for Different Industry Segments

	Brown	Brown w/o ESG	Green	Green w/o ESG
MSE	0.1754	0.1816	0.2165	0.2303
MAE	0.3346	0.3385	0.3692	0.3860

Table: Modelling Performance in Industry Segmentation

Result

R3. Segmentation Model 2 - Industry

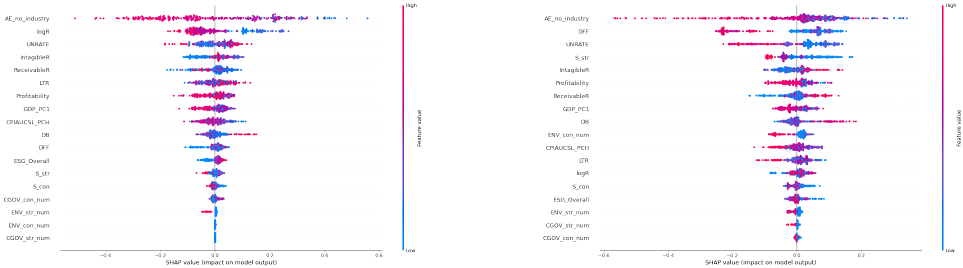


Figure: Summary Plots of SHAP, Brown (Right) vs Green (Left)

Result

R4. Multi-Stage Segmentation

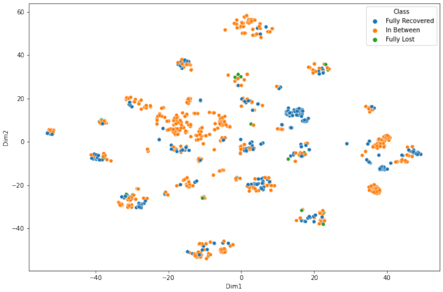


Figure: Spatial Visualization of Three Types of Loss by T-SNE

	Find Fully Lost	Find Fully Recovered	Find In Between
With ESG Variable	**0.6660**	**0.8680**	**0.8338**
Without ESG Variable	0.5794	0.8395	0.8007

Table: How ESG Information Boosts Classification in Multi-Stage Modelling

Result

R5. Rolling Window Regression

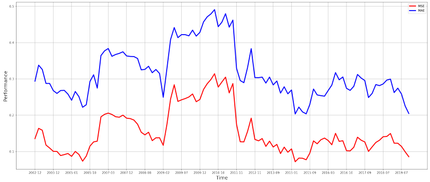


Figure: The Change of Predictive Performance of Selected Variables Over Time

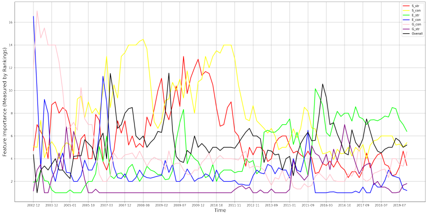


Figure: The Change of Feature Importance of ESG Variables Over Time

Conclusion

- ESG information can enhance the predictive accuracy of LGD estimation as the effective supplement to other proven efficient variable groups and this relationship follows a temporal structure.
- Social perspective play the most significant role in LGD modelling among three pillars of ESG as social perspective includes more information compared with other two dimensions and tends to reflect the risk in the long run.
- ESG information are more effective when estimating LGD of riskier segments.

Future Work

Directions:

- Optimize the modelling framework for LGD
- Seek alternative information source for modelling LGD
- Temporal characteristics of LGD and with PD

Another presentation tomorrow for using information extracted from 10-K to model LGD.

The End

Questions? Comments?