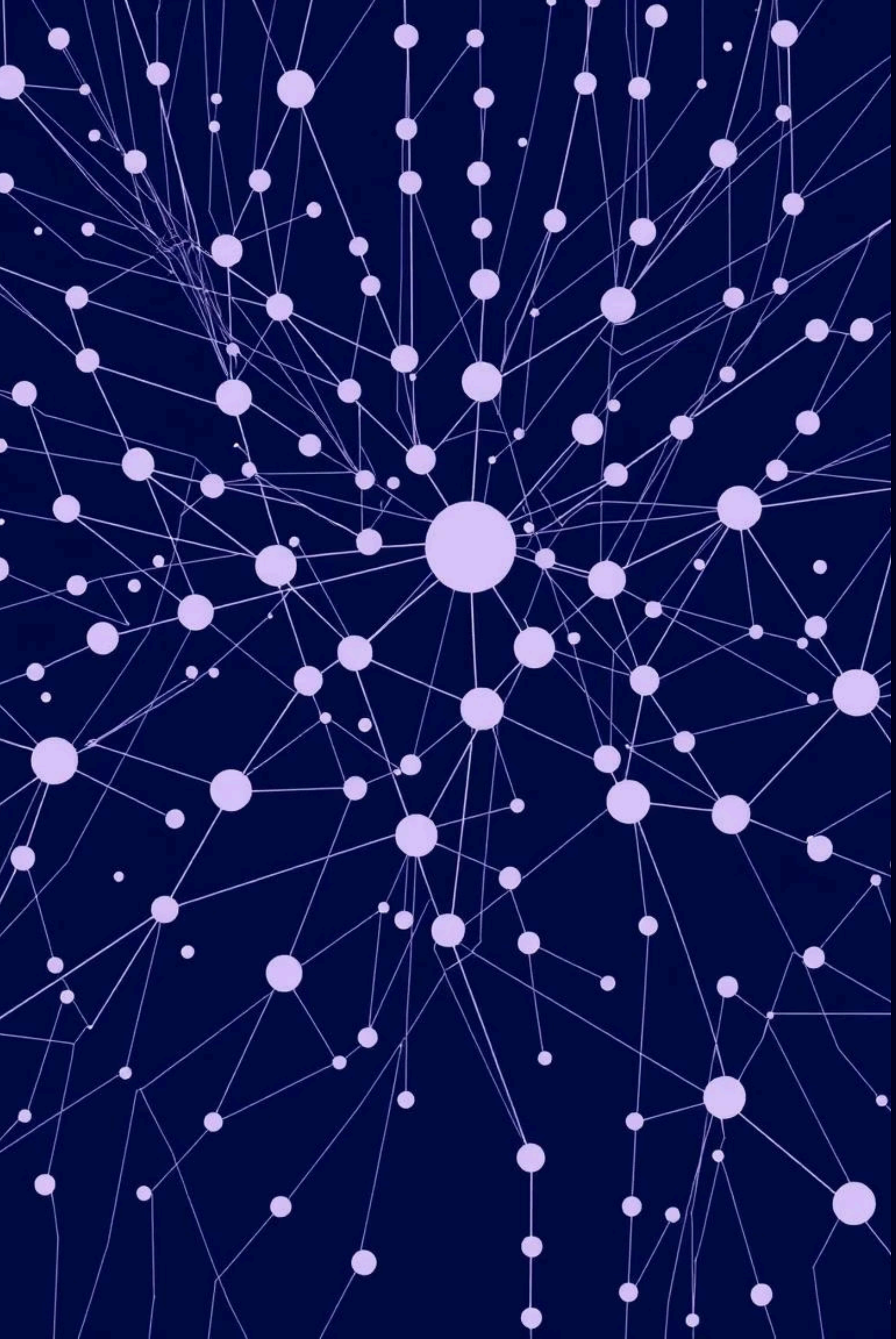


Eco-Risk Alpha

Measuring the Green Discount in Counterparty CVA

Letizia Cazzaniga, Angelique Girod, Marco Piazza, Jacopo Raffaeli

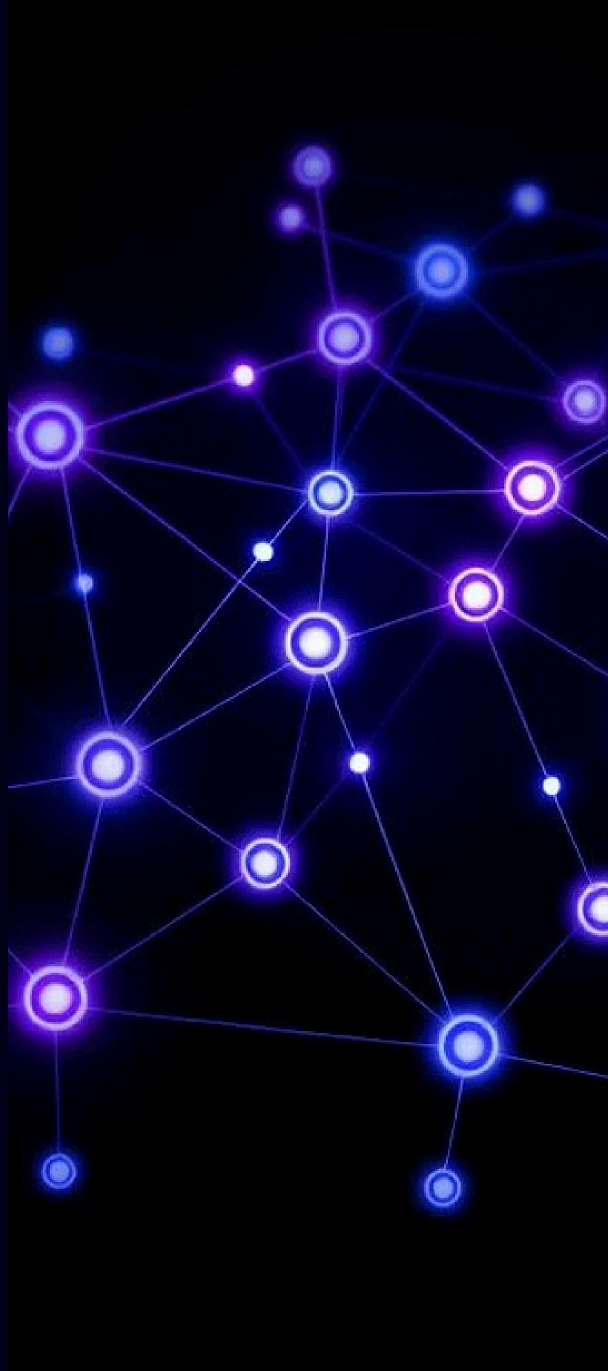
Objective

Predict CDS using Machine Learning

Train Linear and non linear models to explain credit-default-swap (CDS) spreads using ESG and fundamental risk factor variables.

Translate in CVA

Translate those spreads into CVA (Credit Valuation Adjustment) value to quantify a potential ”green credit premium”



Data Analysis Pipeline

Data Exploration and Pre-processing

Preliminary analysis of market data (**CDS spreads**) and risk factors, including **ESG scores** and corporate fundamentals. Data cleaning, missing value management, and normalization to prepare the dataset for modeling.

Model Development and Selection

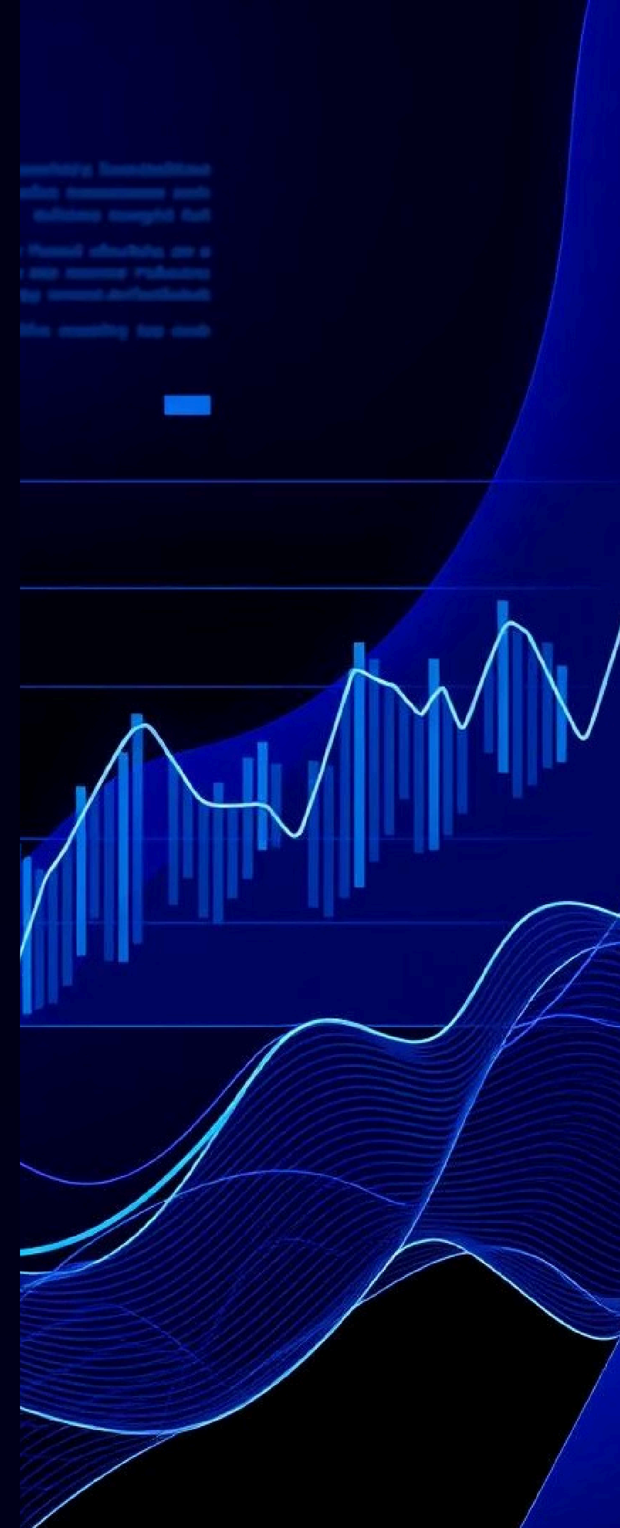
Implementation of **predictive models** (linear and non-linear) to explain CDS spreads. Evaluation of model performance and selection of the most robust framework for integrating ESG risk factors.

CVA Quantification

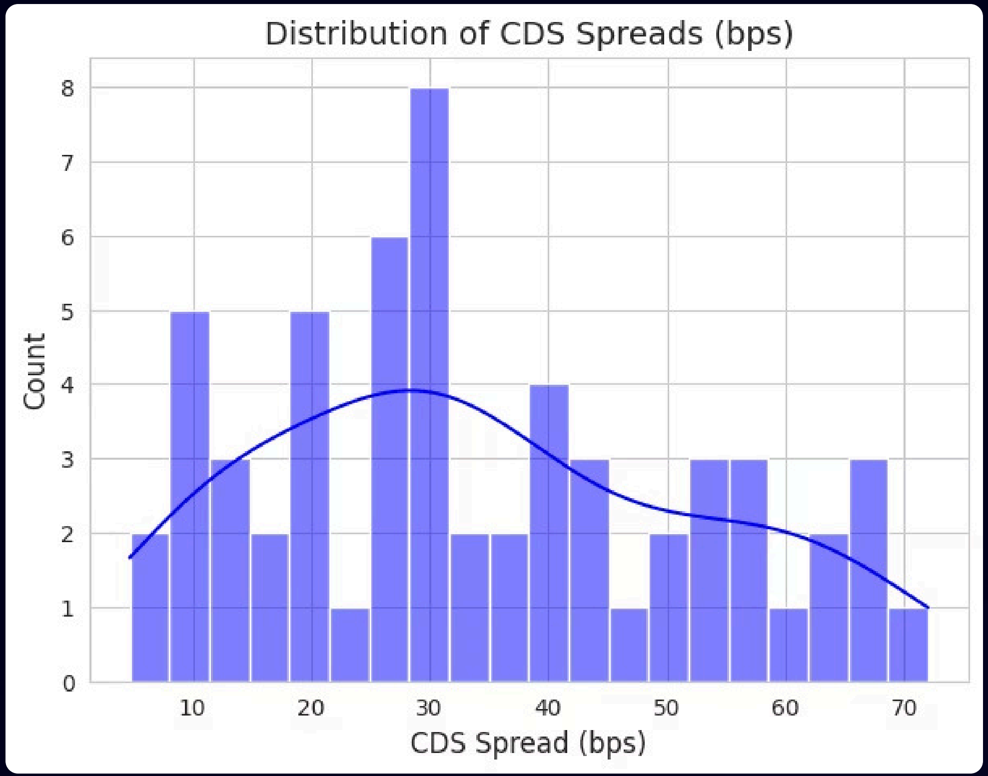
Translation of model outputs into Credit Valuation Adjustment (CVA) to quantify the potential "**Green Credit Premium**," measuring the impact of ESG factors on counterparty default risk.

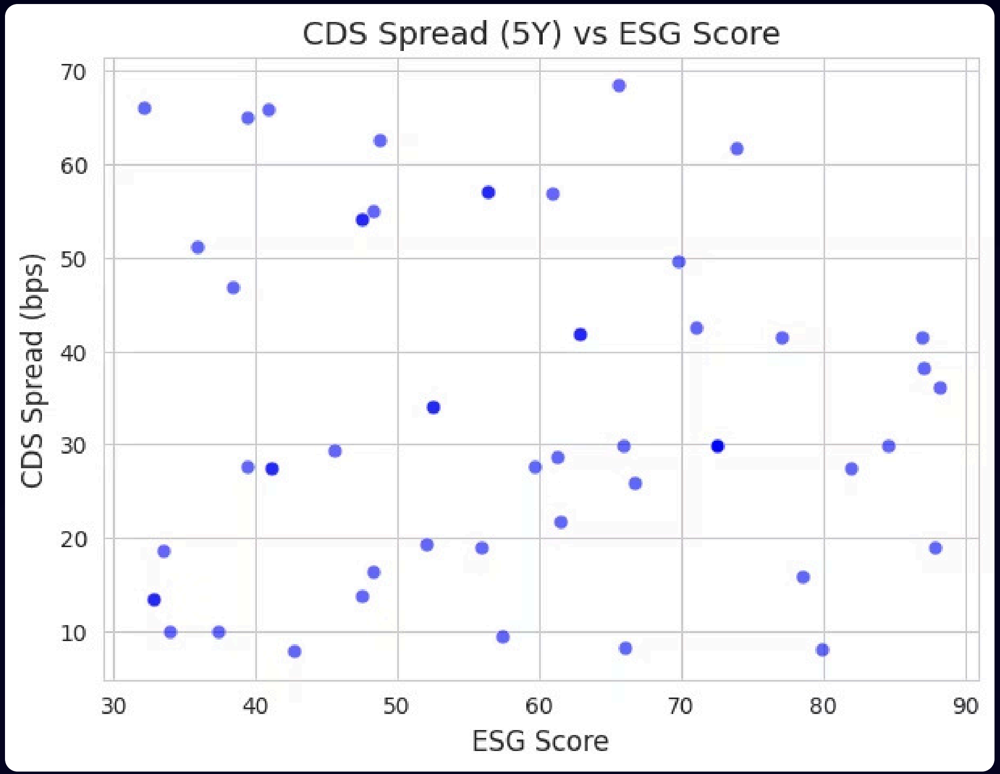
Possible extensions

Extension of the framework simulating new synthetic data and stress test.



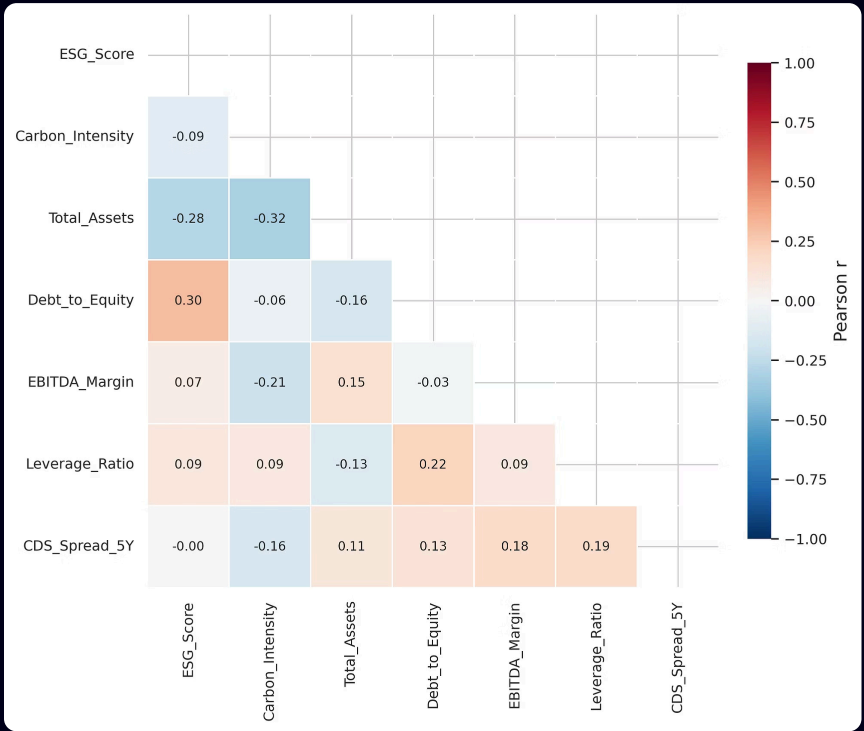
Data Exploration



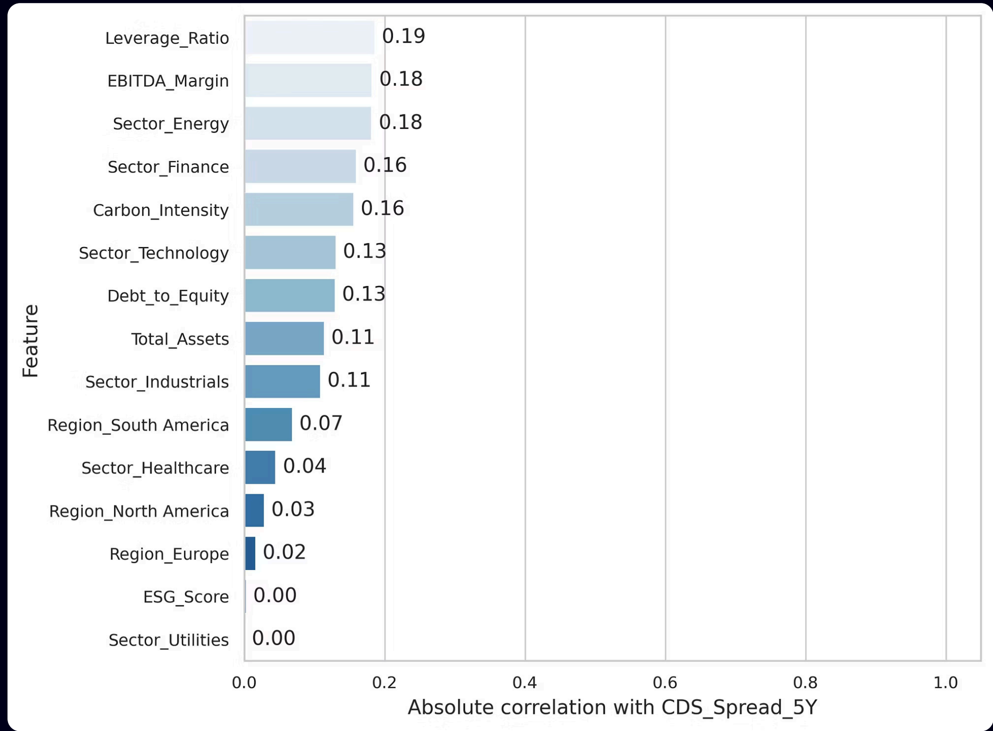


Correlation matrix

Correlation matrix



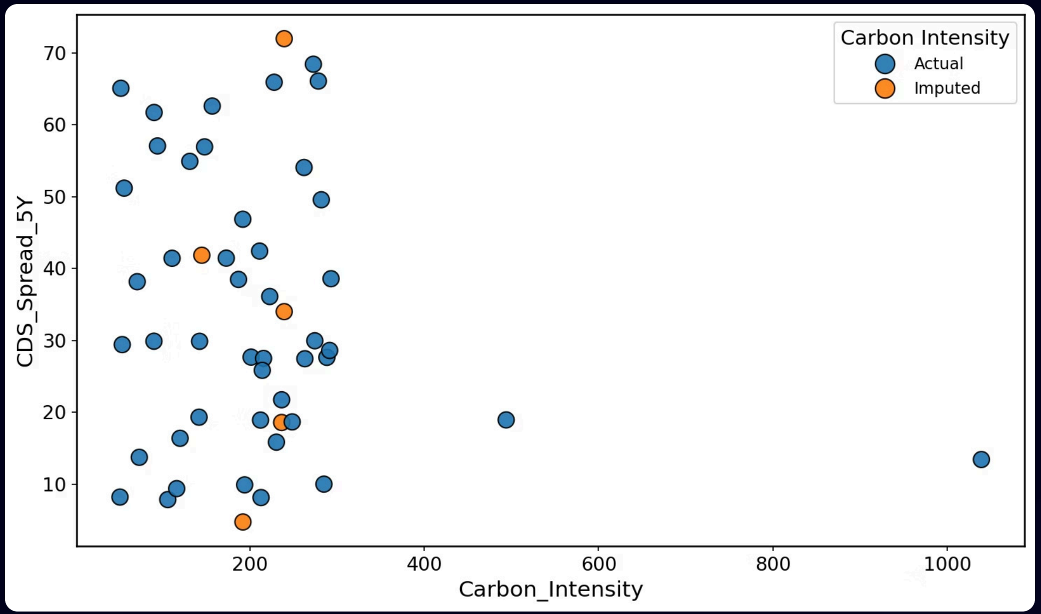
Correlation with target value



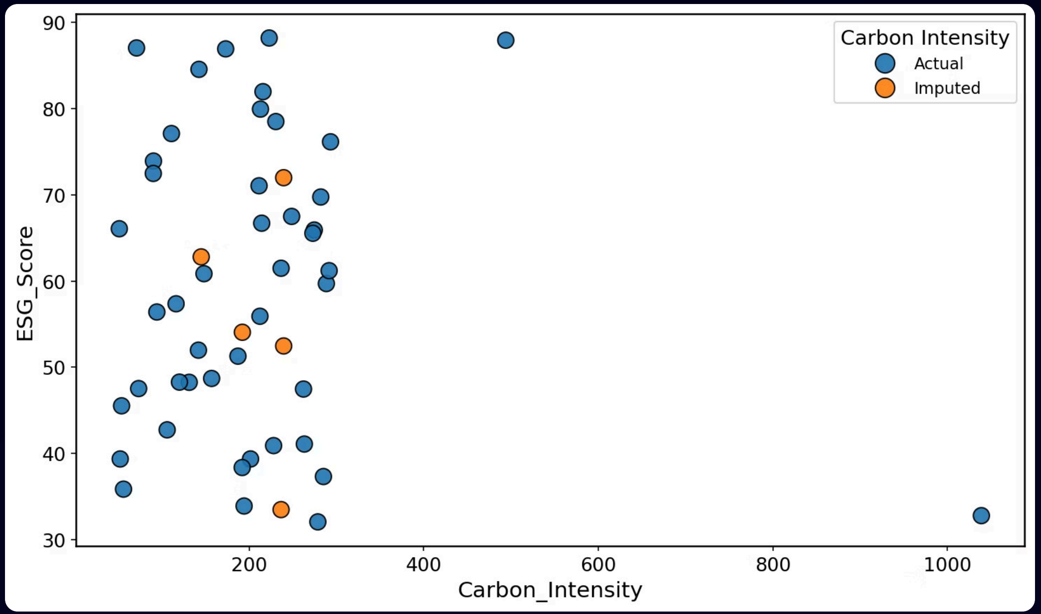
Imputation of Missing Data

Sector-wise median Imputation was performed for missing data imputation

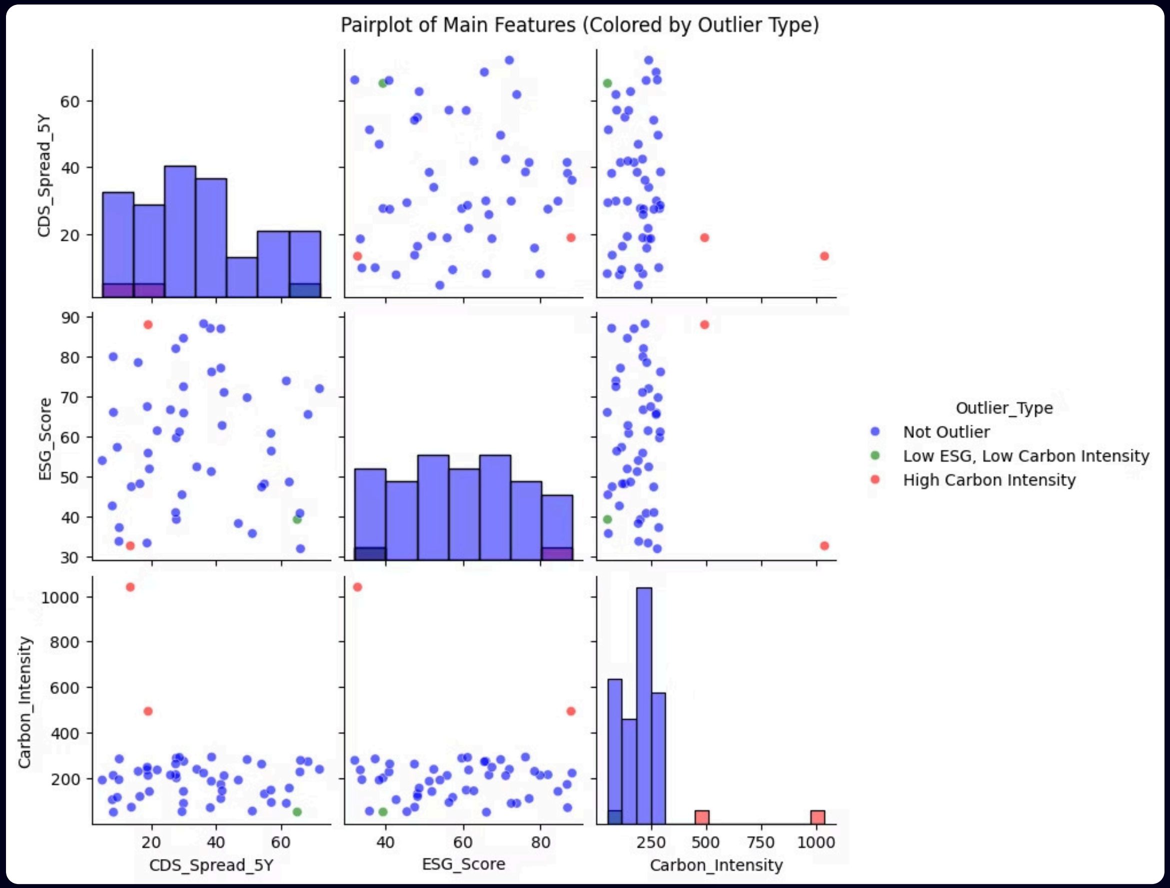
Carbon Intensity vs CDS scores



Carbon Intensity vs ESG scores



Outlier Analysis



- Check of feature value consistency
- Outlier detection using **Isolation Forest**
- Three companies flagged as potential outliers

↓

Maintained to align with the real world

Feature Selection

Sequential Feature Selection (SFS)

Iterative forward and backward

Correlation

Correlation between features and target

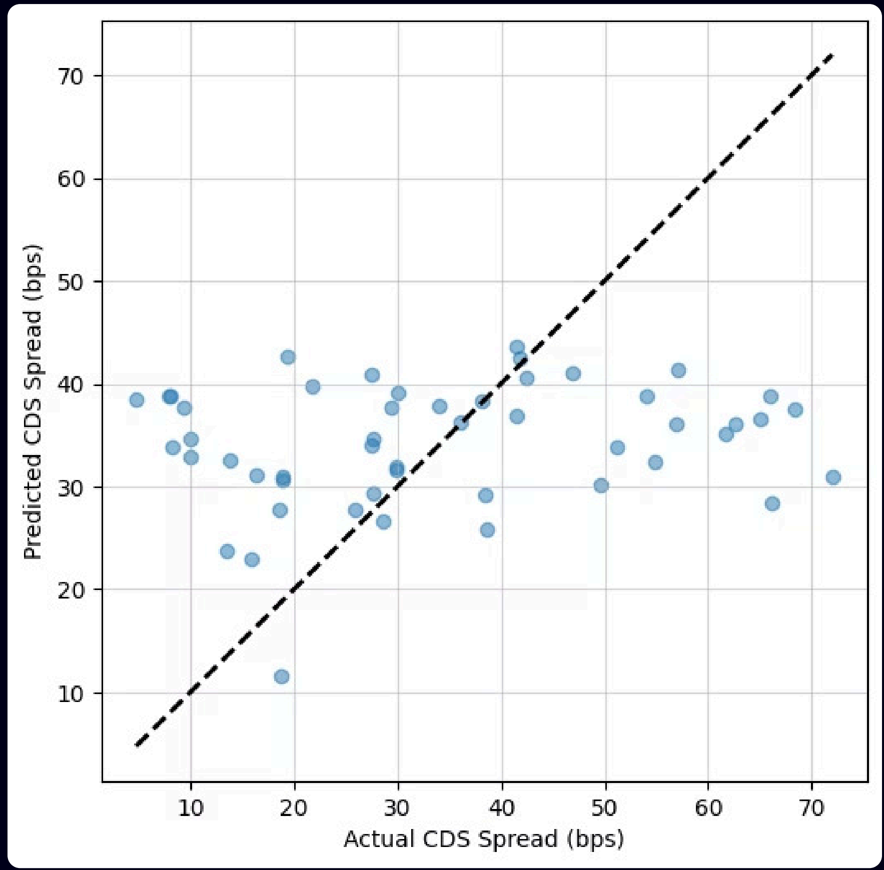
Expert Selection

Manual domain-based choice

Linear Models

Feature Selection	R^2 (Train)	RMSE (CV)
LR full dataset	**0.2**	24.75
LR with Lasso	0	19.3
SFS (Forward and Backward)	0.11	**18.74**
Correlation with target (N=3)	0.1	19.39

SFS Forward / Backward

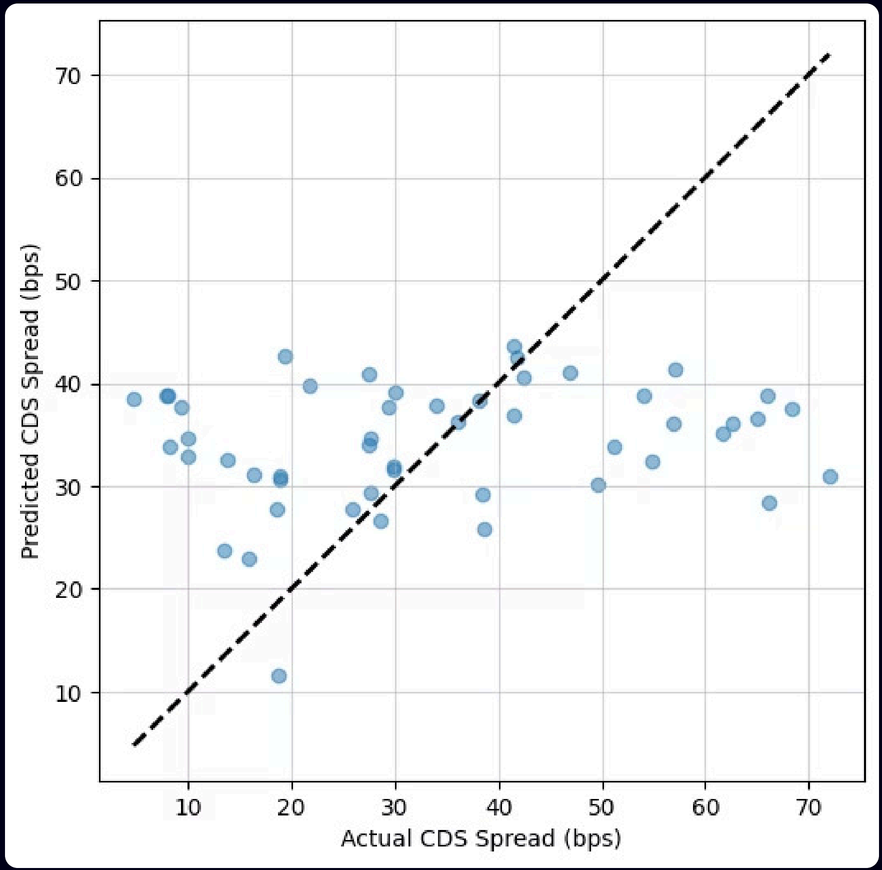


***evaluated on a Leave-One-Out procedure*

Linear Models

Feature Selection	R^2 (Train)	RMSE (CV)
LR full dataset	**0.2**	24.75
LR with Lasso	0	19.3
SFS (Forward and Backward)	0.11	**18.74**
Correlation with target (N=3)	0.1	19.39

SFS Forward / Backward



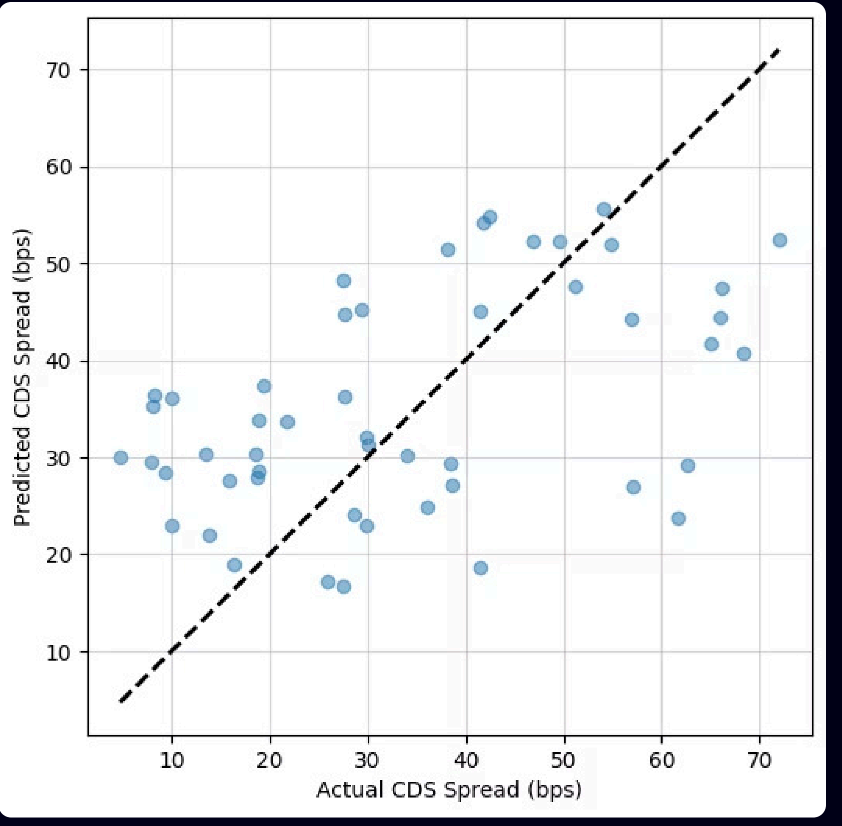
Features selected: Carbon Intensity, Leverage Ratio, Sector Energy

***evaluated on a Leave-One-Out procedure*

Random Forest Model

Feature Selection	R^2 (Train)	RMSE (Test)
All features	0.84	20.44
All features (tuned hyperparameters)	0.77	19.68
SFS Forward	0.88	18.03
SFS Backward	**0.89**	17.1
SFS Forward (tuned hyperparameters)	0.77	17.48
SFS Backward (tuned hyperparameters)	0.78	**16.76**

SFS Backward (tuned hyperparameters)

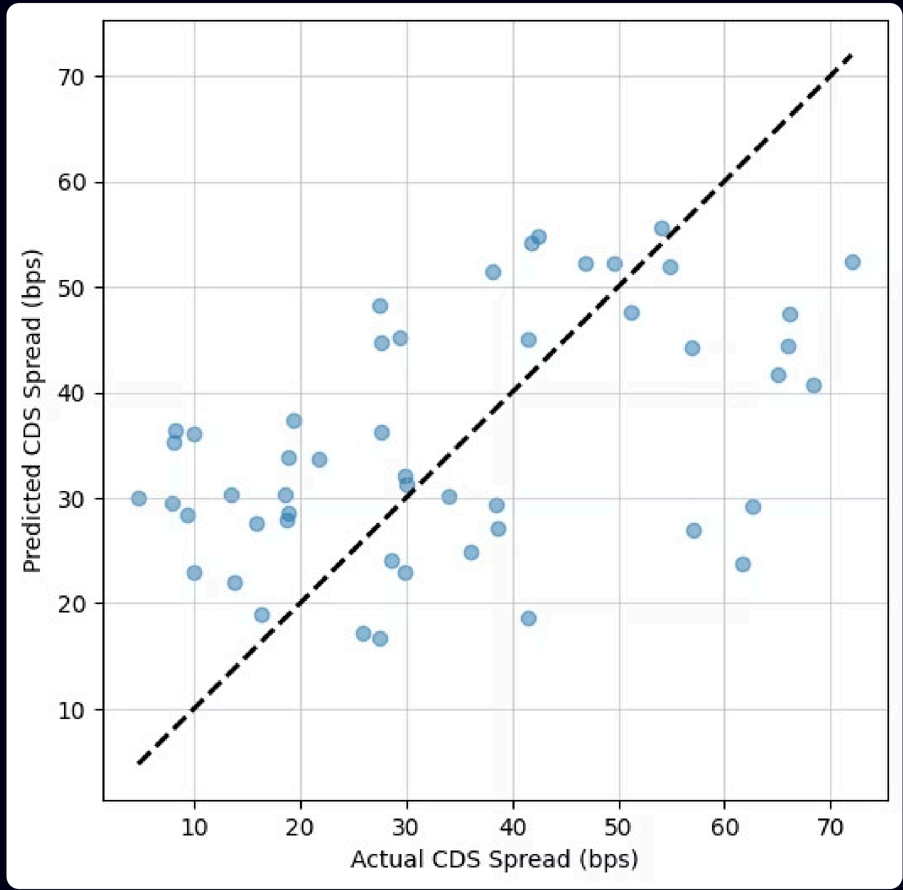


***evaluated on a Leave-One-Out procedure*

Random Forest Model

Feature Selection	R^2 (Train)	RMSE (Test)
All features	0.84	20.44
All features (tuned hyperparameters)	0.77	19.68
SFS Forward	0.88	18.03
SFS Backward	**0.89**	17.1
SFS Forward (tuned hyperparameters)	0.77	17.48
SFS Backward (tuned hyperparameters)	0.78	**16.76**

SFS Backward (tuned hyperparameters)

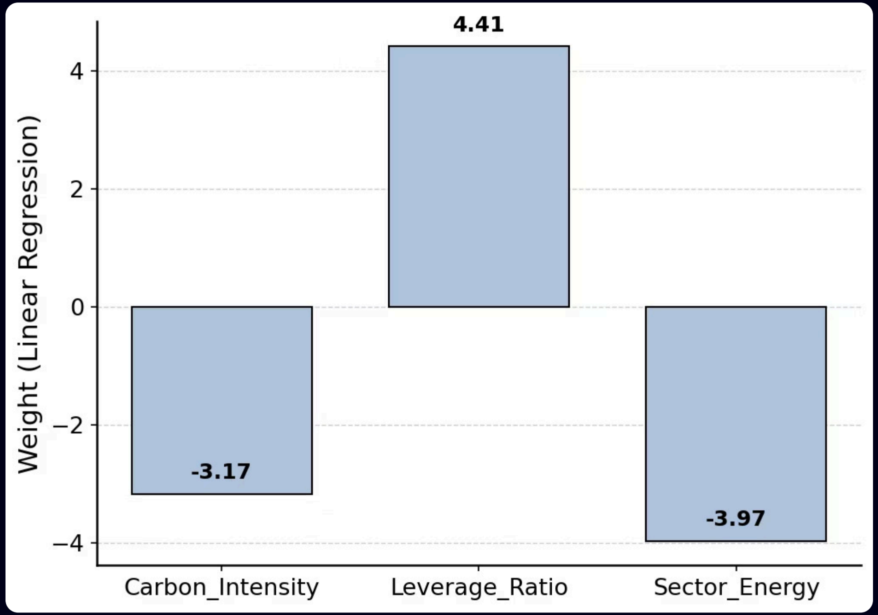


Features Selected: EBITDA Margin, Leverage_Ratio, Sector Energy

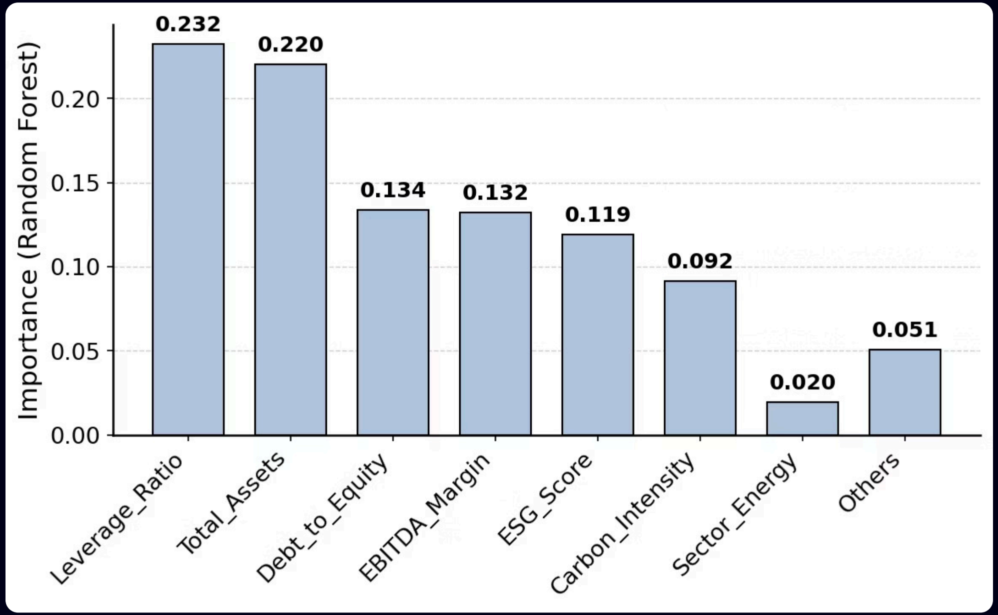
***evaluated on a Leave-One-Out procedure*

Interpretability

Linear Regression



Random Forest



Insights and Quantification of the "Green Premium"

CVA Comparison

Compare actual CVA with that predicted by models.

ESG categorization

Group companies by ESG tercile for a clear comparative analysis.

Green Premium Analysis

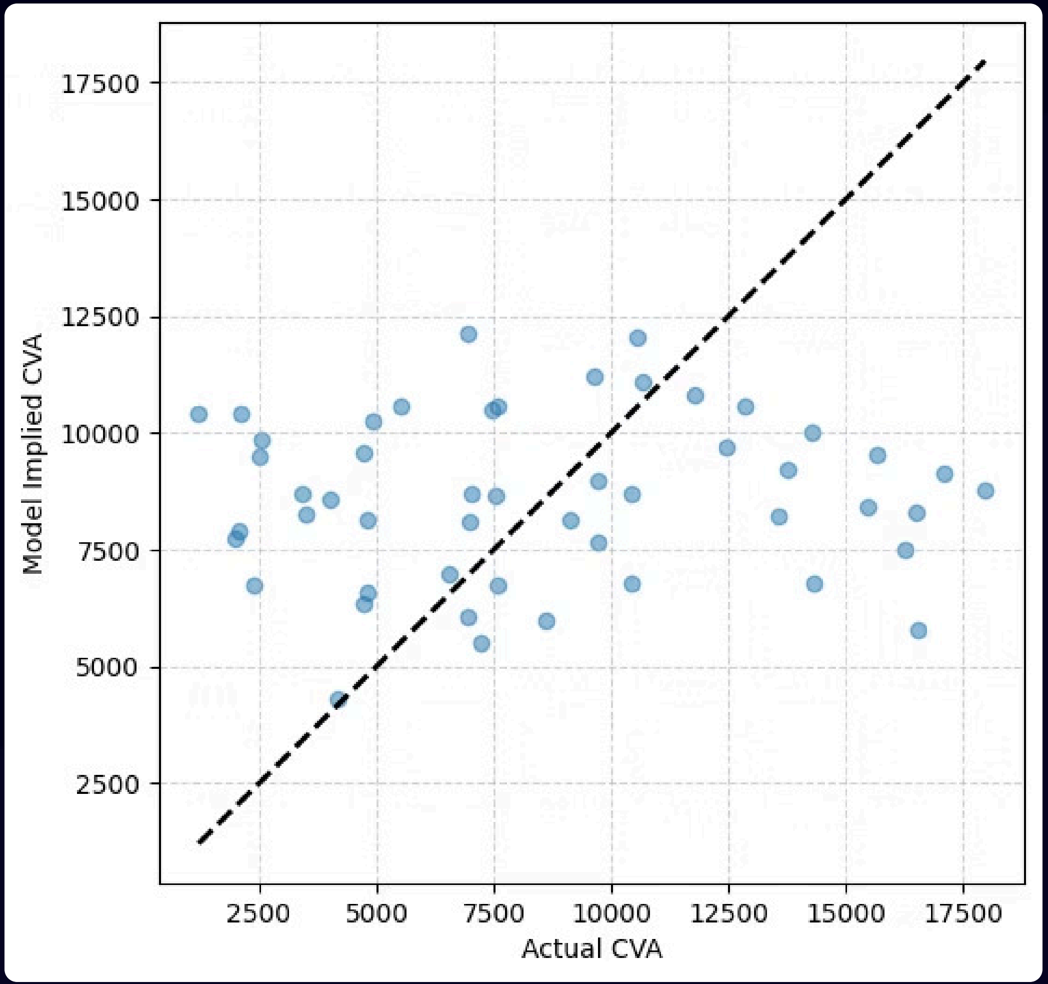
Distribution of Green Premium and interpretion of the economic significance of the green premium

CVA Calculation

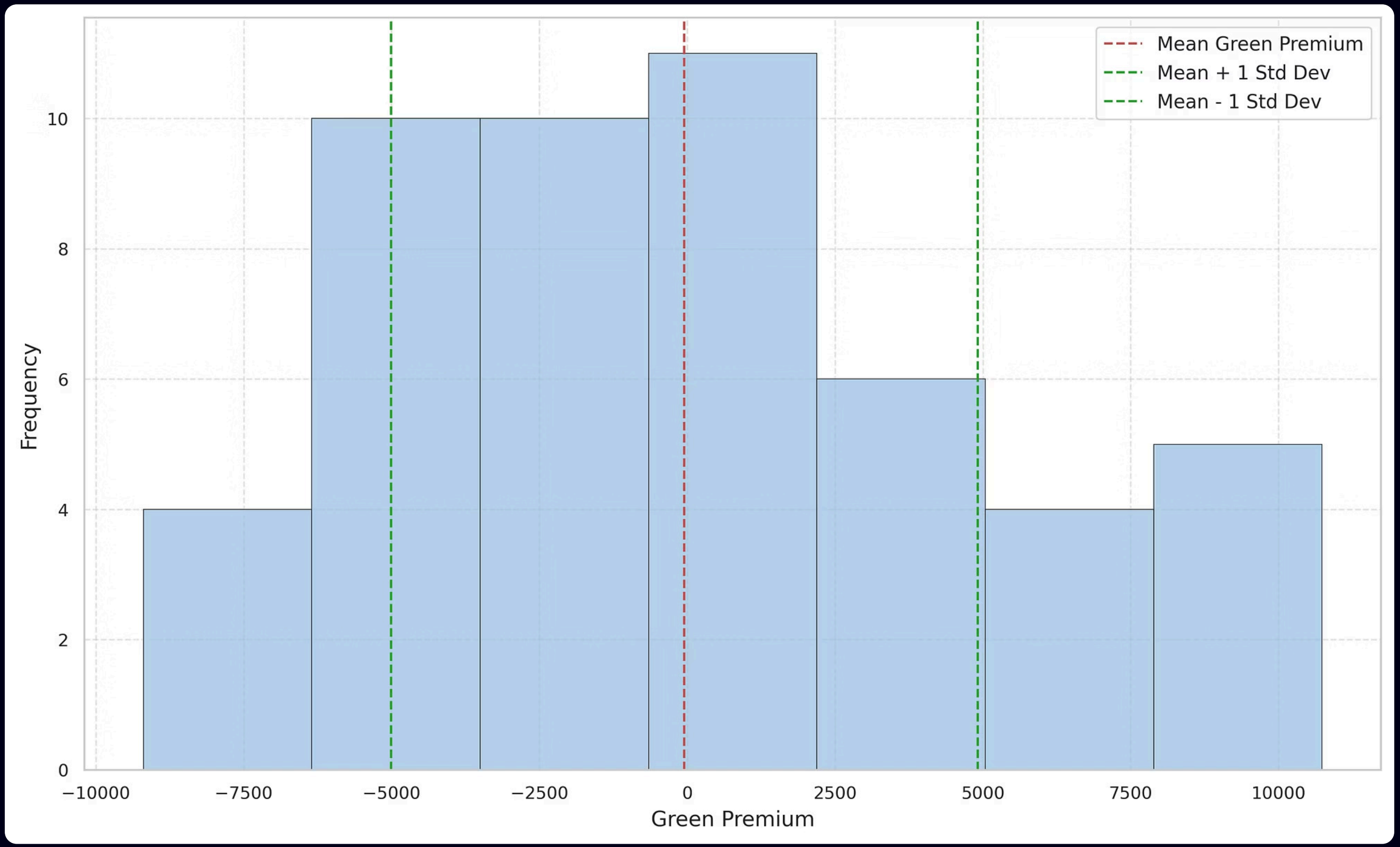
- **CDS Spread to Hazard Rate Conversion:** Transform CDS spreads, both market and model-predicted, into default rates (hazard rates).
- **CVA Formula Application:** Use the unilateral CVA formula, applying it to the provided Expected Exposure (EE) profile. The assumption of a fixed Loss-Given-Default (LGD) of 0.6 is a key parameter in this calculation.

$$CVA \approx LGD \cdot \sum_{i=1}^{N} EE_i \cdot DF_i \cdot (S_{i-1} - S_i)$$

- $S_i = \exp(-\lambda \cdot t_i)$
- $\lambda = \frac{CDS_{Spread}}{1-R}$



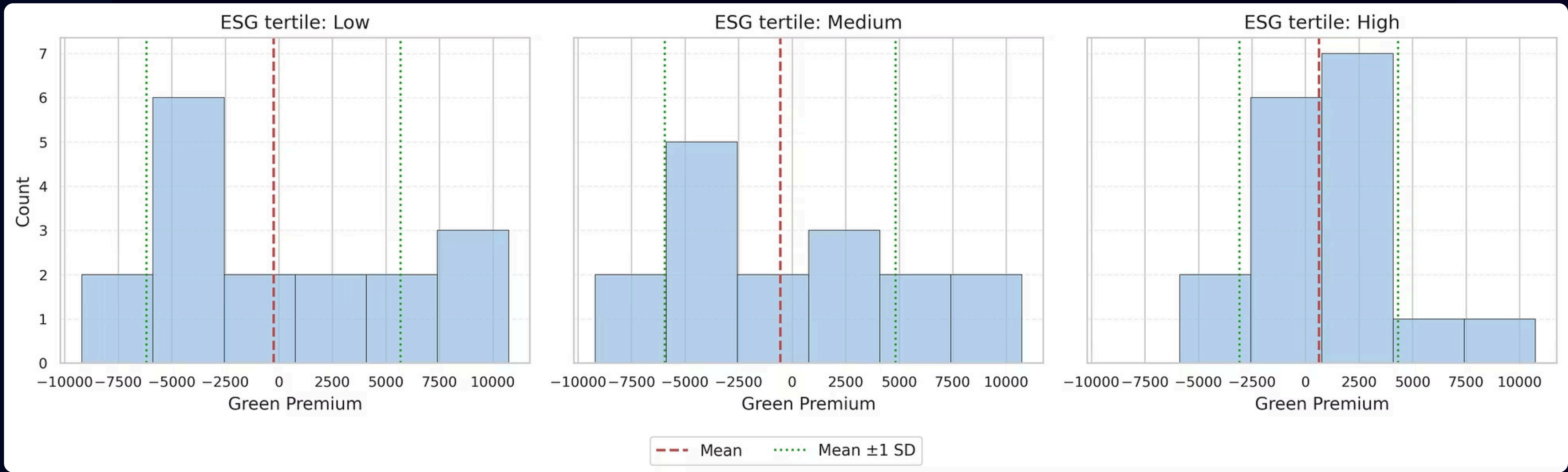
Green Premium Analysis



$$Green\,Premium = CVA_{actual} - CVA_{predicted}$$

Green Premium Analysis

Green premium bucket firms by ESG tertile



Conclusions

- **Objective**

  Test whether ESG factors contribute to explaining CDS spreads, asses whether this implies greener firms benefit from a “green premium” in terms of lower CVA charges.

- **What the model does**

  Our modeling framework can highlight whether ESG factors have statistical relevance for CDS spreads.

  Any claim beyond this would require a broader empirical strategy

- **What the model does not**

  The difference between actual and model-predicted CDS (and CVA) is simply the model error.

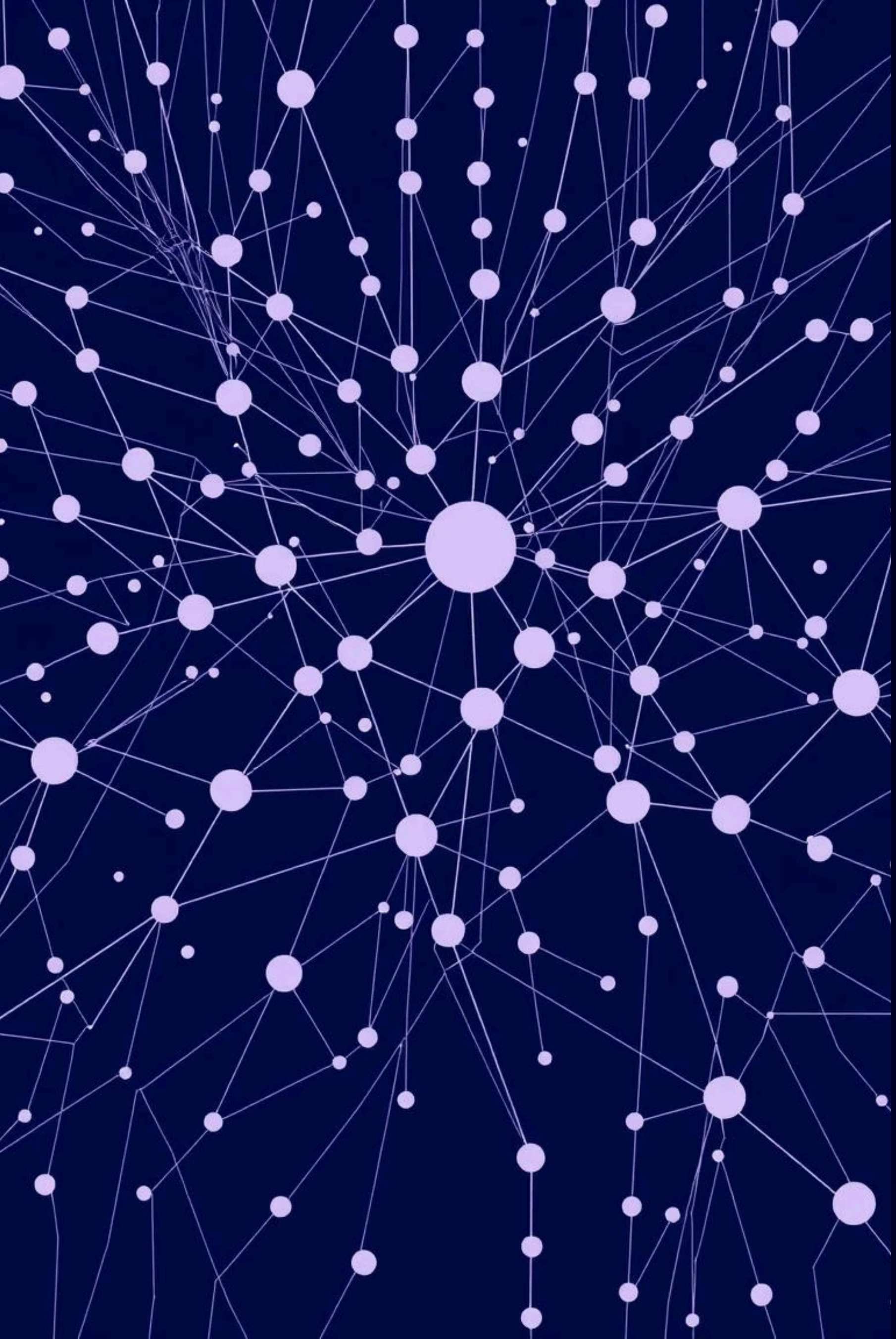
  This means that the gap we label as a “green premium” is not in itself evidence of a true pricing differential in credit markets.

Limitations & Future Developments

- ## Scarcity of data
  - Only ~ 50 observations available
  - Same dataset used for tuning & testing → reliability issues
  - CV helps but is not a substitute for a true test set or nested CV

- ## Synthetic data generation
  - Small sample size makes synthetic data generation unreliable
  - Artificial patterns could distort results more than help

- ## Stress testing
  - No stress tests performed on the zero rate curve
  - Important next step to assess CVA sensitivity under different scenarios



Thank you for your attention!

Letizia Cazzaniga, Angelique Girod, Marco Piazza, Jacopo Raffaeli