

Stock Return

Data:

- 1811 data points

2014 - 2021

Independent variables:

- Beta
- E Score
- S Score
- G Score
- lnMKcap
- InP/B

Target:

- Annualized Stock
Return Next Year

$$Stock\ Return(t + 1) = \sum_{k=1}^m \beta_k \times X_k(t)$$

	ret+1	e	s	g	lnmkc	lnpb	b
Year							
2021	-15.3142	85.082653	82.101183	80.783313	12.882437	0.311658	1.118366
2021	-25.5563	86.114663	88.772461	61.524124	12.490134	-0.014027	1.405726
2021	-13.9433	82.436049	79.747984	65.274447	11.966134	-0.117137	1.176937
2021	-22.3607	85.869001	73.445377	50.040152	11.108468	0.310239	0.978488
2021	-21.2348	43.553476	75.853462	87.251413	11.061943	0.312388	1.138725

2014	11.5725	18.282112	34.208641	57.583742	7.172459	0.094130	1.204618
2014	-1.7214	18.282112	34.687947	30.876225	7.168756	-0.087168	0.447732
2014	7.5676	18.282112	30.556785	48.110703	6.586729	0.262263	0.252190
2014	-20.1946	17.201540	31.219030	14.858410	7.549217	-0.737808	3.368074
2014	-53.8411	18.282112	35.937658	30.998775	6.773048	-0.209223	1.238775

Stock Return

OLS Regression Results							
Dep. Variable:	ret+1	R-squared:	0.172				
Model:	OLS	Adj. R-squared:	0.168				
Method:	Least Squares	F-statistic:	43.60				
Date:	Tue, 25 Jul 2023	Prob (F-statistic):	1.46e-48				
Time:	13:38:17	Log-Likelihood:	-6048.4				
No. Observations:	1267	AIC:	1.211e+04				
Df Residuals:	1260	BIC:	1.215e+04				
Df Model:	6						
Covariance Type:	nonrobust						
coef	std err	t	P> t	[0.025	0.975]		
const	2.0098	4.800	0.419	0.675	-7.406	11.426	
b	12.7558	2.312	5.518	0.000	8.220	17.291	
e	0.2128	0.068	3.129	0.002	0.079	0.346	
s	-0.0921	0.079	-1.173	0.241	-0.246	0.062	
g	-0.0564	0.047	-1.212	0.226	-0.148	0.035	
lnmkc	-1.6529	0.864	-1.912	0.056	-3.349	0.043	
lnpb	34.1153	2.315	14.736	0.000	29.573	38.657	
Omnibus:	779.268	Durbin-Watson:	1.990				
Prob(Omnibus):	0.000	Jarque-Bera (JB):	18928.913				
Skew:	2.403	Prob(JB):	0.00				
Kurtosis:	21.315	Cond. No.	387.				

Stock Return

Result:

- Ridge Reg. has the best performance
- Optimal $\lambda = 10$
- E Score increases by 1, Stock Return increases by 0.21
- S Score increases by 1, Stock Return decreases by 0.1
- G Score increase by 1, Stock Return decreases by 0.06

Features	model_lin	model_ridge	model_lasso	model_net
0 b	12.755776	11.470183	12.197990	7.155286
1 e	0.212811	0.207920	0.209257	0.193493
2 s	-0.092148	-0.107070	-0.096544	-0.157452
3 g	-0.056381	-0.055327	-0.055713	-0.051708
4 lnmkc	-1.652909	-1.253793	-1.482799	0.056662
5 lnpb	34.115296	31.845153	33.574119	23.665696

	Model	RMSE
0	Linear Regression	25.0809
1	Ridge Regression	25.0404
2	Lasso Regression	25.0628
3	Net Regression	25.1340

Volatility

Data:

- 1811 data points
2014 - 2021

Independent variables:

- Beta
- E Score
- S Score
- G Score
- lnMKcap
- lnP/B

Target:

- Annualized Volatility
Next Year

$$Volatility(t+1) = \sum_{k=1}^m \beta_k \times X_k(t)$$

	v+1	e	s	g	lnmkc	lnpb	b
Year							
2021	29.437980	85.082653	82.101183	80.783313	12.882437	0.311658	1.118366
2021	31.909446	86.114663	88.772461	61.524124	12.490134	-0.014027	1.405726
2021	33.498623	82.436049	79.747984	65.274447	11.966134	-0.117137	1.176937
2021	28.625238	85.869001	73.445377	50.040152	11.108468	0.310239	0.978488
2021	30.123806	43.553476	75.853462	87.251413	11.061943	0.312388	1.138725

2014	19.878259	18.282112	34.208641	57.583742	7.172459	0.094130	1.204618
2014	14.606244	18.282112	34.687947	30.876225	7.168756	-0.087168	0.447732
2014	14.255208	18.282112	30.556785	48.110703	6.586729	0.262263	0.252190
2014	49.486026	17.201540	31.219030	14.858410	7.549217	-0.737808	3.368074
2014	82.740428	18.282112	35.937658	30.998775	6.773048	-0.209223	1.238775

Volatility

OLS Regression Results						
Dep. Variable:	v+1	R-squared:	0.211			
Model:	OLS	Adj. R-squared:	0.207			
Method:	Least Squares	F-statistic:	56.14			
Date:	Tue, 25 Jul 2023	Prob (F-statistic):	1.35e-61			
Time:	13:52:57	Log-Likelihood:	-5218.9			
No. Observations:	1267	AIC:	1.045e+04			
Df Residuals:	1260	BIC:	1.049e+04			
Df Model:	6					
Covariance Type:	nonrobust					
coef	std err	t	P> t	[0.025	0.975]	
const	40.9181	2.494	16.407	0.000	36.025	45.811
e	-0.0113	0.035	-0.319	0.750	-0.081	0.058
s	0.0763	0.041	1.870	0.062	-0.004	0.156
g	-0.0212	0.024	-0.876	0.381	-0.069	0.026
lnmkc	-3.2324	0.449	-7.196	0.000	-4.114	-2.351
lnpb	-8.2872	1.203	-6.889	0.000	-10.647	-5.927
b	16.7872	1.201	13.975	0.000	14.430	19.144
Omnibus:	348.706	Durbin-Watson:	1.922			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	835.966			
Skew:	1.487	Prob(JB):	2.97e-182			
Kurtosis:	5.645	Cond. No.	387.			

Volatility

Result:

- Elastic Net Reg. has the best performance
- Optimal $\lambda = 0.0954$ Optimal ratio = 0.5
- E Score increases by 1, Volatility increases by 0.08
- S Score increases by 1, Volatility decreases by 0.02
- G Score increase by 1, Volatility decreases by 3

	Features	model_lin	model_ridge	model_lasso	model_net
0	b	16.787159	15.859723	16.684675	15.151033
1	e	-0.011280	-0.013155	-0.011677	-0.014778
2	s	0.076319	0.076899	0.076161	0.077193
3	g	-0.021175	-0.019528	-0.020942	-0.018213
4	Inmkc	-3.232419	-3.099725	-3.212385	-2.993789
5	Inpb	-8.287166	-8.052948	-8.218156	-7.816660

	Model	RMSE
0	Linear Regression	14.6462
1	Ridge Regression	14.6178
2	Lasso Regression	14.6432
3	Net Regression	14.6028

Boosting Methods

We tried XGBoost regression and LightGBM Regression. However, both of the tree-based models have lower model R_squared and higher RMSE compared to linear regression.

	Model	R_squared	RMSE
Stock Return Model	XGBoost	-0.023	27.75
	LightGBM	0.017	27.19
Volatility Model	XGBoost	-3.116	32.32
	LightGBM	0.023	15.75

Volatility

Beta-ESG Models

Objective

Predicting the effect of E, S, and G pillar scores on the following year's Beta

$$\beta_{T+1} = \sum_{k=1}^m d1_k * \beta_{T,k} + d2_k * E_{T,k} + d3_k * S_{T,k} + d4_k * G_{T,k}$$

Equ. 1

Data Selection

- 300 U.S. banks' Beta, E, S, and G pillar score from the year 2014 to 2021
- Banks with >30% null were dropped
- Remaining null data entries imputed with column mean
- Scaled data using preprocessing.normalize()
- **End Total = 221 banks with 7 years of E, S, G pillar score and 8 years of Beta**

Models

1. Year by Year OLS model grouped by MarketCap
2. Overall Model including all years' data (Beta + RoA)



Beta-ESG Model Finding

Year By Year model

All banks (221)

The significance of the S and G pillar scores shows improvement. The result aligns with the hypothesis that ESG scores gain more and more attention as an investment evaluation factor.

Large and Mid-Large banks (6)

The results show no significance of all 3 ESG pillar scores. ESG scores are overlooked by the market due to the information overload of larger banks. However, the results might be invalid due to lack of data.

Medium banks (110)

The significance of the Social pillar scores shows clear improvement, with only 2020's data being an outlier. Although G scores show improvement, their p-value fluctuates between years.

Small banks (105)

The Social and Governance scores show no significance for most years while showing some improvement in recent years. This may be caused by missing data in the early on.

**Large and Mid-Large banks (6) model has only 6 data points each yrs
**All results have high R-squared, with the lowest at 0.726

Beta-ESG Model Finding

Overall model 1

$$\beta_{T+1} = \sum_{k=1}^m d1_k * \beta_{T,k} + d2_k * E_{T,k} + d3_k * S_{T,k} + d4_k * G_{T,k}$$

Equ. 1

	Observations	R^2	Environmental		Social		Governance	
			Coef.	P-value	Coef.	P-value	Coef.	P-value
All banks (221 banks)	1547	0.920	-0.0342	0.001	0.1791	0.000	0.1094	0.000
Large & Mid-large (6 banks)	42	0.987	-0.1122	0.202	0.6000	0.013	-0.0697	0.429
Medium (110 Banks)	770	0.933	-0.0270	0.051	0.2067	0.000	0.1391	0.000
Small (105 Banks)	735	0.902	-0.0039	0.783	0.1647	0.000	0.0903	0.002

Table 1

Beta-ESG Model Finding

Overall model 2

$$\beta_{T+1} = \sum_{k=1}^m d1_k * ROA_{T,k} + d2_k * E_{T,k} + d3_k * S_{T,k} + d4_k * G_{T,k}$$

Equ. 2

	Observations	R^2	Environmental		Social		Governance	
			Coef.	P-value	Coef.	P-value	Coef.	P-value
All banks (210 banks)	1470	0.847	-0.0718	0.000	0.4452	0.000	0.3469	0.000
Large & Mid-large (6 banks)	42	0.981	-0.1843	0.093	1.1906	0.000	-0.0228	0.022
Medium (111 Banks)	777	0.893	-0.0573	0.001	0.4722	0.000	0.3710	0.000
Small (93 Banks)	651	0.782	0.0246	0.280	0.3195	0.000	0.3039	0.000

Table 2

Single Variable Input Regression (SVIR)

Models

Objective

Determine which individual components of E, S, and G contribute the most to the overall pillar score

Data Selection

- 136 independent variables from list of Refinitiv's most important banking ESG components
- Independent variables with >50% null dropped
- Variables with <50% missing imputed with column mean
- **End Total = 106 Independent Variables**

Models

1. OLS Regression
2. Classification & Regression Tree (CART)
3. Random Forest



SVIR Findings

OLS REGRESSION

ENVIRONMENTAL

- Adj. R-Squared = 0.967
- 11 Coefficients w/ statistical significance

Top 5 Most Important Features

1. Assets Under Management
2. Renewable Products
3. Equator Principles
4. Environmental Products
5. Environmental Restoration Initiatives

SOCIAL

- Adj. R-Squared = 0.993
- 17 Coefficients w/ statistical significance

Top 5 Most Important Features

1. Product Access Low Price
2. Corporate Responsibility Awards
3. Policy Diversity Opportunity
4. Day Care Services
5. Fundamental Human Rights ILO UN

GOVERNANCE

- Adj. R-Squared = 0.87
- 8 Coefficients w/ statistical significance

Top 5 Most Important Features

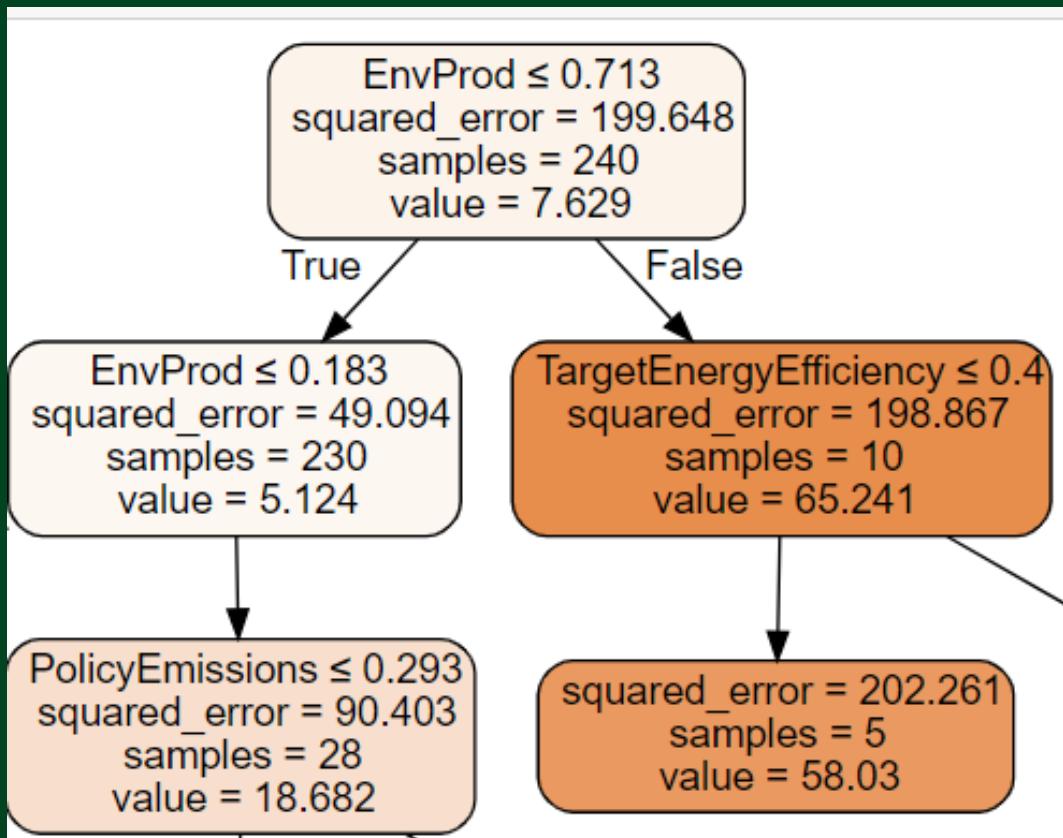
1. Shareholder Rights Policy (-)
2. CSR Sustainability Reporting
3. CEO Chairman Duality
4. Board Structure Policy
5. Global Compact Signatory

SVIR Findings

CART

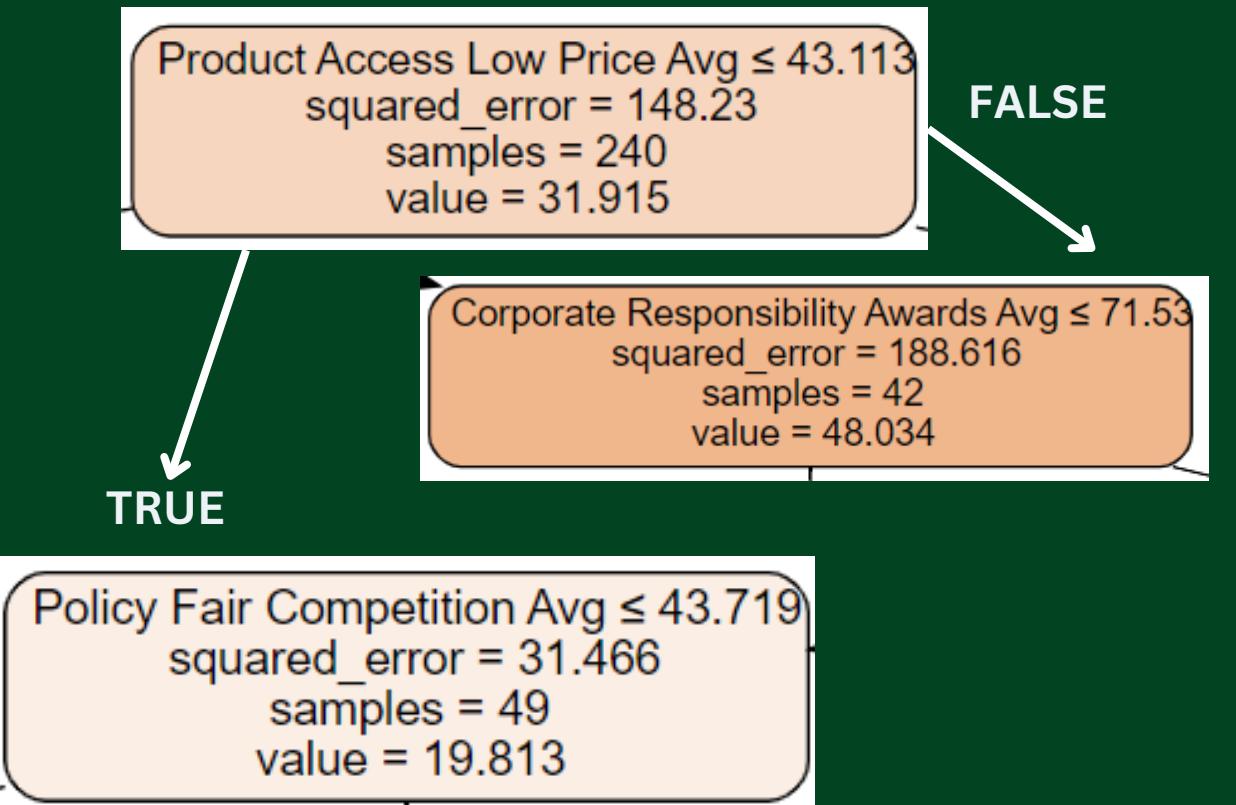
ENVIRONMENTAL

- MSE = 71.14
- R-Squared = 0.32
- Adj. R-Squared = -0.08



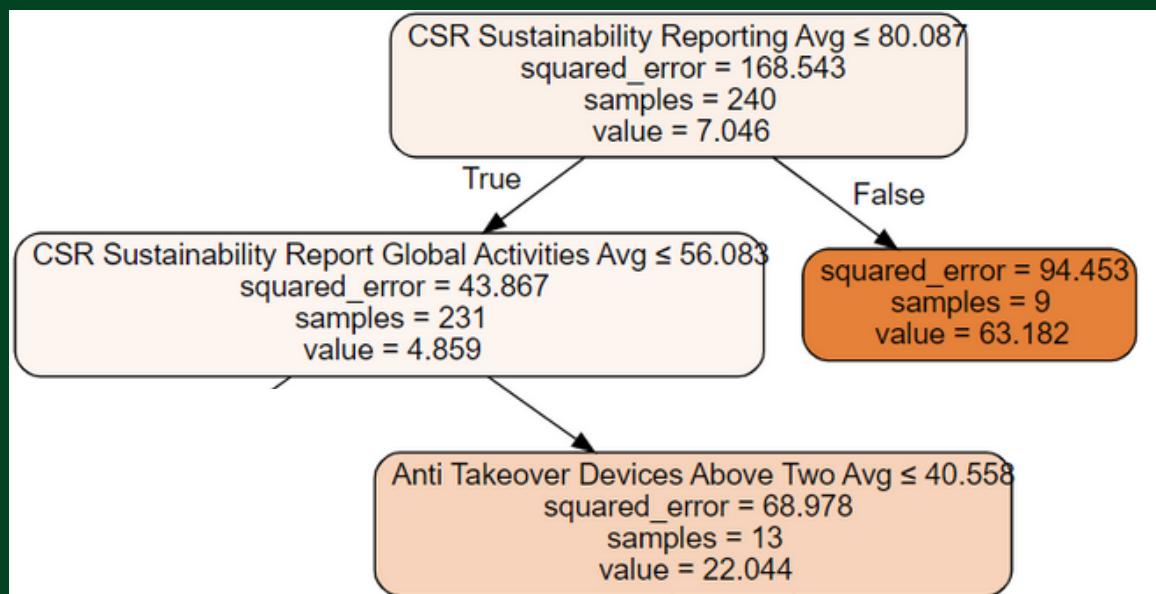
SOCIAL

- MSE = 29.83
- R-squared = 0.63
- Adj. R-Squared = 0.22



GOVERNANCE

- MSE = 60.22
- R-squared = 0.30
- Adj. R-Squared = -5.86



SVIR Findings

RANDOM FOREST

ENVIRONMENTAL

- Adj. R-Squared = 0.85
- MSE = 16.21

Top 5 Most Important Features

1. **Environmental Products**
2. **Equator Principles**
3. Energy Efficiency
4. **Renewable Products**
5. Targets Emissions

SOCIAL

- Adj. R-Squared = 0.86
- MSE = 11.52

Top 5 Most Important Features

1. **Corporate Responsibility Awards**
2. **Product Access Low Price**
3. Policy Diversity Opportunity
4. Policy Fair Competition
5. Whistleblower Protection

GOVERNANCE

- Adj. R-Squared = 0.62
- MSE = 33.17

Top 5 Most Important Features

1. **CSR Sustainability Reporting**
2. CSR Sustainability Report Global Activities
3. CSR Sustainability External Audit
4. CSR Sustainability Committee
5. Highest Renumeration Package

GRANULARITY MODELS

Attempt to make use of the massive quantity of data available in the Refinitiv database.



Multiple Linear Regression

- ScikitLearn's Ridge Regression Algorithm
- Benefit: the ability to monitor independent variables that are highly correlated



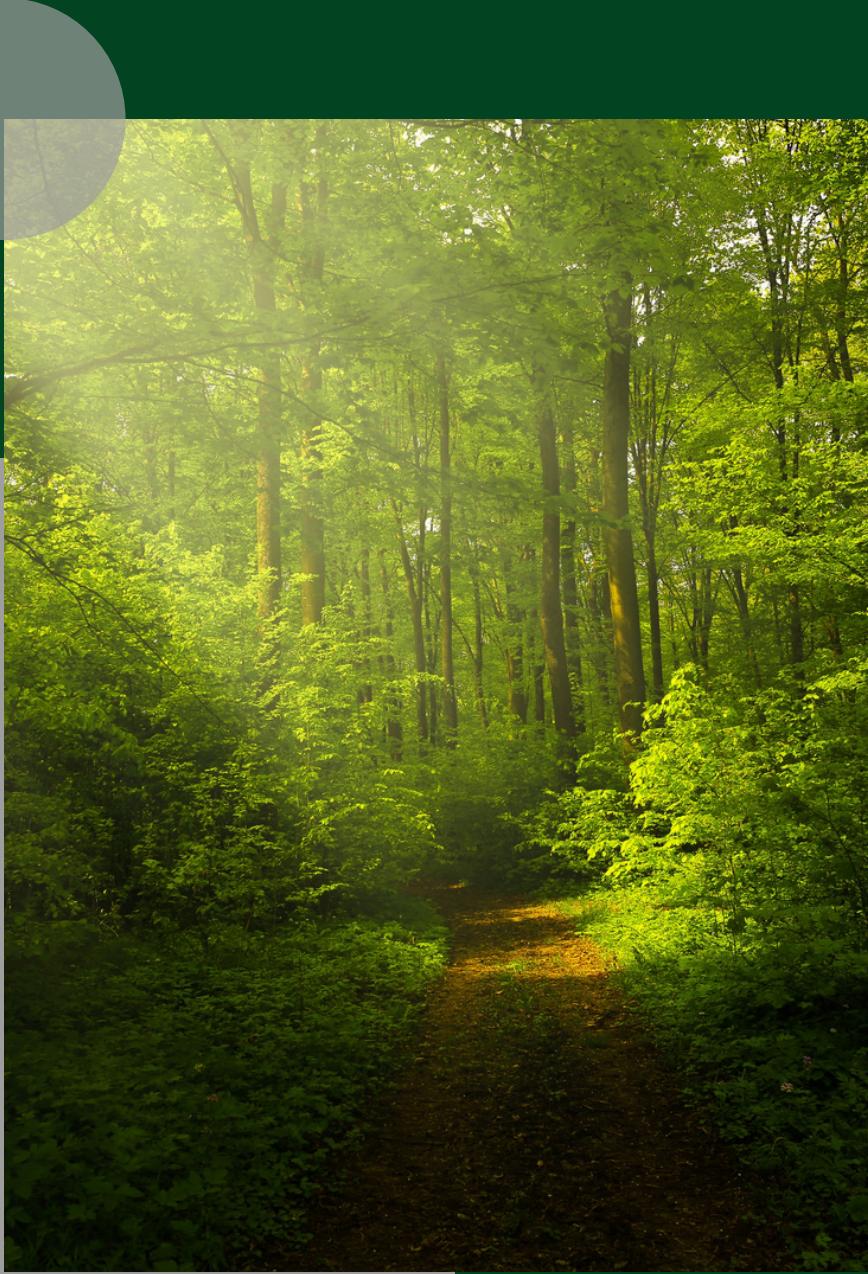
Ordinary Least Squared

- Statmodels
- Benefit: allowed the regression models to be better fitted



Support Vector Regression (SVR)

- ScikitLearn's SVR Algorithm
- Benefit: to best predict outputs for correlations



ESG V PERFORMANCE

DATA SPECS

10

Years of ESG Scores

244

Total Features by Year

70%

Completion Threshold

84

ESG Features for P/B Models

10

Years of Correlation

80

ESG Features for Volatility and Stock Return Models

Target variables were bank performance KPIs: Stock Returns, Price to Book, and Volatility

Independent ESG variables were determined by correlation analysis to determine which fields held value.



ESG V BANK PERFORMANCE

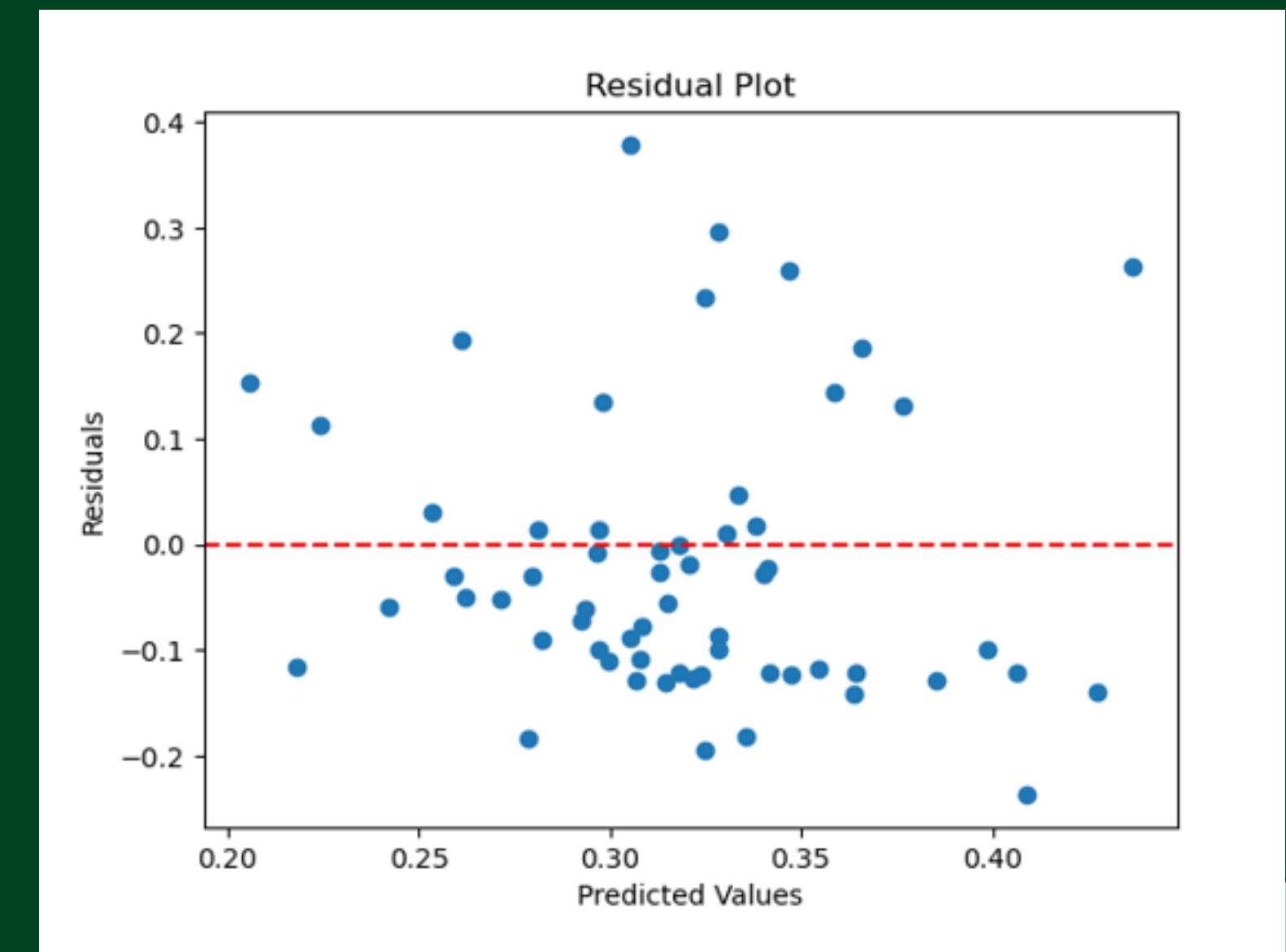
SUPPORT VECTOR MODEL RESULTS

Target: Price to Book

Mean Mean Absolute Error (MAE):
0.079087623080858

Mean R²: 0.1621802346089828

The Support Vector model was by far the worst and mostly was used to contrast and understand the difficulty in drawing meaningful natural distinctions in the data.



SAVE ENVIRONMENT

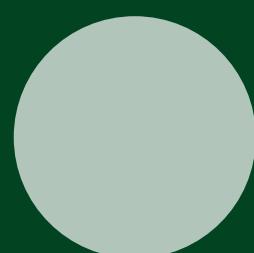
RESULTS FROM RIDGE MODEL

FY 2022

Target: Volatility

Mean MAE: 0.056540042619307694

Mean R^2: 0.06905844309390427



Feature	Coefficient
Policy Fair Competition	0.037475
Independent Board Members	-0.141032
Anti Takeover Device Above Two	-0.007580
Management Score	0.039947
Average Board Tenure	-0.085835
Shareholders Score	-0.026904
Board Gender Diversity	0.037088
Executive Members Gender Diversity	0.003654
Product Responsibility Score	0.022124



RESULTS OLS REGRESSION

Target: Volatility
FY 2022

Mean MAE: 0.05600178937483312
Mean R^2: 0.09538010398033211

The best models came out of an OLS model
p-value maximum reduction algorithm.

OLS Regression Results						
Dep. Variable:	Volatility	R-squared:	0.109			
Model:	OLS	Adj. R-squared:	0.081			
Method:	Least Squares	F-statistic:	3.942			
Date:	Sun, 30 Jul 2023	Prob (F-statistic):	9.42e-05			
Time:	21:41:48	Log-Likelihood:	309.17			
No. Observations:	300	AIC:	-598.3			
Df Residuals:	290	BIC:	-561.3			
Df Model:	9					
Covariance Type:	nonrobust					
		coef	std err	t	P> t	[0.025 0.975]
const		0.3202	0.051	6.312	0.000	0.220 0.420
Policy Fair Competition		0.0321	0.012	2.772	0.006	0.009 0.055
Independent Board Members		-0.1913	0.062	-3.097	0.002	-0.313 -0.070
Anti Takeover Devices Above Two		-0.0015	0.028	-0.055	0.956	-0.056 0.053
Management Score		0.0467	0.024	1.964	0.050	-8.92e-05 0.094
Shareholders Score		-0.0039	0.021	-0.183	0.855	-0.046 0.038
Average Board Tenure		-0.0617	0.031	-1.987	0.048	-0.123 -0.001
Executive Members Gender Diversity, Percent		-0.0194	0.028	-0.698	0.486	-0.074 0.035
Board Gender Diversity, Percent		0.0306	0.027	1.136	0.257	-0.022 0.084
Product Responsibility Score		0.0392	0.024	1.663	0.097	-0.007 0.086
Omnibus:	245.778	Durbin-Watson:	2.078			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	6225.167			
Skew:	3.111	Prob(JB):	0.00			
Kurtosis:	24.431	Cond. No.	29.7			

ESG V PERFORMANCE

RESULTS OLS REGRESSION

Target: Stock Return

FY 2016

Mean MAE: 0.05963211613264909

Mean R^2: 0.48338119024015

Most of the models generated negative R^2 values, something relatively rare in data modeling, and reserved for models that were exceptionally awful at predicting the target feature. Even those models that were predictive were not consistent across all 9 or 10 years.

		coef	std err	t	P> t	[0.025
	0.975]					
const		0.7683	0.113	6.808	0.000	0.545
0.992						
Announced Layoffs To Total Employees		0.5279	0.128	4.113	0.000	0.273
0.782						
Equal Shareholder Rights Score		0.1411	0.035	4.074	0.000	0.072
0.210						
Corporate Responsibility Awards Score		-0.1132	0.026	-4.311	0.000	-0.165
-0.061						
Director Election Majority Requirement		0.0682	0.022	3.100	0.002	0.025
0.112						
Board Meeting Attendance Average		-0.0932	0.033	-2.855	0.005	-0.158
-0.029						
Board Background and Skills Score		-0.1565	0.045	-3.448	0.001	-0.246
-0.067						
Shareholders Score		-0.1798	0.043	-4.191	0.000	-0.265
-0.095						
Sustainability Compensation Incentives Score		0.0828	0.031	2.682	0.008	0.022
0.144						
Targets Diversity and Opportunity Score		-0.2056	0.053	-3.873	0.000	-0.311
-0.100						
Audit Committee Mgt Independence Score		0.0900	0.034	2.658	0.009	0.023
0.157						
Shareholder Rights Policy Score		-0.2635	0.104	-2.526	0.013	-0.470
-0.057						
Anti Takeover Devices Above Two		-0.1538	0.048	-3.211	0.002	-0.249
-0.059						
Quality Mgt Systems Score		-0.4395	0.090	-4.862	0.000	-0.619
-0.260						
Total Senior Executives Compensation To Revenues in million		-0.1878	0.044	-4.296	0.000	-0.275
-0.101						
Internal Promotion Score		0.0793	0.033	2.409	0.018	0.014
0.145						
Nomination Committee Independence Score		-0.0796	0.024	-3.306	0.001	-0.127
-0.032						
Board Specific Skills, Percent		-0.1205	0.045	-2.672	0.009	-0.210
-0.031						
Board Attendance Score		0.2579	0.093	2.783	0.006	0.074
0.442						
Auditor Tenure		-0.1179	0.047	-2.516	0.013	-0.211
-0.025						
Public Availability Corporate Statutes Score		0.1999	0.063	3.161	0.002	0.075
0.325						
Human Rights Contractor Score		-0.2022	0.064	-3.153	0.002	-0.329
-0.075						

Key Takeaways:



Across all models, Price to Book was the most consistent and easier to predict

No model was truly and consistently predictive, with occasional swings on certain years

In general, even specific ESG data and values tend to be, by themselves, terrible predictors of future financial success

OUR DELIVERABLES



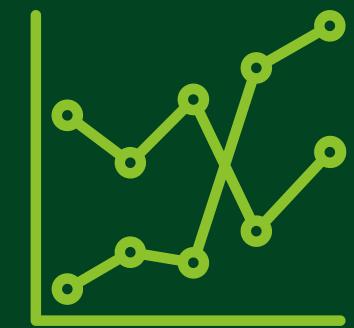
**PROJECT
REPORT**



**FINAL
PRESENTATION**



**DATA
FILES**



**JUPYTER
NOTEBOOK**

QUESTION & ANSWER

Any further explanations we can provide?

More detailed explanations and processes can also be found in the written report and Jupyter notebooks.

Were we able to add value to your exploration of ESG?

We hope our models' findings can be useful to your team moving forward.

What are your initial thoughts on our work?

Are there any areas of concern? Any additional models you would have liked to see?

Are the findings what you expected?

We knew the ESG data would provide limitations from the start. Did anything surprise you?

THANK YOU

LOYOLA UNIVERSITY MSBDA CAPSTONE

Deloitte.



Preparing people to lead extraordinary lives