UNIVERSITÀ CATTOLICA del Sacro Cuore

DIPARTIMENTO DI SCIENZE ECONOMICHE E SOCIALI

GET RID OF THE SCORE! ESG RISK AND SUSTAINABLE FINANCE

Gianni Guastella
Matteo Mazzarano
Stefano Pareglio

Quaderno n. 163/settembre 2024

VP VITA E PENSIERO

Università Cattolica del Sacro Cuore

DIPARTIMENTO DI SCIENZE ECONOMICHE E SOCIALI

GET RID OF THE SCORE! ESG RISK AND SUSTAINABLE FINANCE

Gianni Guastella
Matteo Mazzarano
Stefano Pareglio

Quaderno n. 163/settembre 2024

VP VITA E PENSIERO

Gianni Guastella, Dipartimento di Matematica e Fisica, Università Cattolica del Sacro Cuore, Brescia.

Matteo Mazzarano, Dipartimento di Scienze Politiche e Internazionali, Università degli Studi di Siena.

Stefano Pareglio, Dipartimento di Matematica e Fisica, Università Cattolica del Sacro Cuore, Brescia.

giovanni.guastella@unicatt.it

matteo.mazzarano@unisi.it

stefano.pareglio@unicatt.it

I quaderni possono essere richiesti a:
Dipartimento di Scienze Economiche e Sociali,
Università Cattolica del Sacro Cuore
Via Emilia Parmense 84 - 29122 Piacenza - Tel. 0523 599 342
http://dipartimenti.unicatt.it/dises

dises-pc@unicatt.it

www.vitaepensiero.it

ISBN 978-88-343-5869-6

Abstract. Despite the pressures on financial institutions for ESG integration into portfolio strategies, evidence suggests this is often confined to ESG screening. This paper confronts ESG screening with an integration strategy leveraging raw ESG data to construct portfolios with the lowest risk. Matching Eikon ESG data with Merton's distance to default (DtD), the paper compares the internal and external validity of models estimated using the aggregated ESG score and the raw data, applying regularisation techniques in the latter case. Using raw data leads to an ESG portfolio with low risk and high returns compared to the standard aggregate ESG screening approach.

Keywords. Financial risk, ESG Score, sustainable finance, LASSO regression.
J.E.L. **classification.** Q56, G15, G24, G30

G Guastella acknowledges financial support from PNRR MUR project ECS_00000033_ECOSISTER

1. Introduction

Financial institutions' attention to ESG (acronym for Economic, Social, and Environmental) factors has considerably gained momentum riding the wave of the 2015 Paris Agreement (PA). Pressured by governments, supervisory authorities, and consumers, financial institutions started their sustainability journey by integrating ESG considerations into strategic decision-making (Amel-Zadeh & Serafeim, 2018). According to a recent European Banking Authority report (Coleton et al., 2020), almost 95% of financial institutions that responded to a survey on ESG integration reported that sustainability is incorporated into the business strategy.

ESG metrics represent a core tool of this paradigmatic change from ”Sustainable Finance 1.0”, characterised by sin stock avoidance and the prevalence of financial valuation over ESG aspects, to ”Sustainable Finance 2.0”, which recognises ESG factors as a source of long-term value creation (Schoenmaker & Schramade, 2018). Specialised providers aggregate a large amount of information publicly disclosed by the graded companies to produce simple metrics -scores or grades- that are easy to interpret and comparable across companies, making it easier for investors to design ESG-aligned strategies.

In practice, ESG metrics are far from reaching their designed objectives. Billio et al. (2021) find substantial heterogeneity in the characteristics, attributes, and standards of the E, S, and G components across the ratings of different providers. Berg et al. (2022) compare the data of six score providers and confirm a substantial divergence, a significant part of which can be ascribed to how different providers measure the same ESG factors. The 2020 OECD Business and Finance Outlook (OECD, 2020) also suggests a low correlation of ESG scores with E scores and, most importantly, environmental and climate metrics.

Three are the drawbacks of using aggregate ESG scores in portfolio strategies. Firstly, ESG scores measure disclosure more than performance. Raw information is sourced from non-financial disclosures, news, public statements, and company websites and includes actual performance metrics and disclosure metrics essentially based on boolean indicators. The data quality varies across categories (Ehlers et al., 2024) because while performance metrics follow specific protocols (like GHG emissions), disclosure metrics do not.

Secondly, scores are subject to the problem of additive aggregations (ECJRC and OECD, 2008). A company may excel in some categories and fail in others. As a result, companies with the same grade may represent very different investment risks and opportunities (Erhart, 2022). Finally, ESG scores are produced in a manner that is not necessarily consistent with the purpose of their use. Provider's objective is to return a synthetic indicator representing heterogeneous information about companies' ESG risks and opportunities. These stem from higher profitability (Luo & Bhattacharya, 2006), reduced information asymmetry (Cho et al., 2013) and lower financing costs (Cheng et al., 2014), among others. While a correlation between financial performance and a few specific ESG factors is expected, that correlation may weaken or disappear when considering the overall ESG score (Guastella et al., 2022).

Alternatively, raw ESG data can directly inform investment strategies after identifying the ESG factors relevant to the investor's preferences and objectives. Machine Learning (ML) techniques allow the selection of relevant information within a set based on each piece of information's explanatory power and have already been applied to ESG mining (Lanza et al., 2020; Margot et al., 2021; De Lucia et al., 2020). These studies suggest that portfolios identified with ML approaches outperform those created with the more conventional ESG screening.

More recent literature has shifted the ESG focus from financial performance to credit risk, arguing that a better ESG disclosure reduces information asymmetry and cash-flow volatility and is associated with lower Distance-to-Default (DtD) and tail risk (Atif & Ali, 2021; Bax et al., 2023). ML applications in this field are limited to Bonacorsi et al. (2022), which shows that ESG raw data can predict creditworthiness but does not compare this approach with the ESG screening.

This paper applies linear model regularisation (James et al., 2021), particularly LASSO regression, to select the most influential ESG factors determining credit risk variation across European listed companies. Results are used to construct an ad hoc ESG indicator called "ESGRisk" that considers only relevant ESG information. The external validity of the LASSO approach is assessed, demonstrating that it has a better out-of-sample predictive capacity than the standard ESG score or the E-S-G single-category scores. Additionally, the financial performance of two portfolios constructed using the ESGRisk and ESG scores are compared, showing that ESGRisk outperforms the ESG score.

The paper's contribution is twofold. Firstly, it brings new evidence on ESG impact assessment by looking at the company's riskiness, which is less explored in literature, notwithstanding the regulatory authorities' concerns (EBA, 2020). Secondly, it provides evidence that ESG raw data have a greater potential for informing portfolio strategies than aggregated raw scores (Ehlers et al., 2024). In light of ESG screening limitations (Berg et al., 2022), this paper proposes a strategy to detect only relevant ESG factors and create a tailored ESG risk score to be integrated into the institution's risk management framework.

The remainder of the paper includes a methodology section (2), the empirical results (3), and a validation section (4). A discussion with implications for decision-makers concludes the work.

2. *Methodology and data*

2.1 Methodology

The empirical analysis starts from the simple model in Equation 1 to estimate the impact of ESG factors on risk using the reduced form in Atif & Ali (2021), linking the company-level measure of risk (DtD_i) to its determinants, among which the ESG scores (ESG_i), and the literature-based control variables (X_i). ESG is included either:

1. as an aggregated score,
2. as disaggregated E-S-G pillar scores,

3. as a set of single categories (e.g. the “Emissions score”), or
4. as a vector of raw data.

$$y_i = \alpha + \beta \cdot ESG_i + \gamma ' X_i + \varepsilon_i \quad (1)$$

Using the full vector of raw data leads to large *K* (number of variables) over *N* (number of companies), and a linear model regularisation approach (James et al., 2023, Ch 6) is used to prevent overfitting and improve predictions. More specifically, the Least Absolute Shrinkage and Selection Operator (LASSO) approach is employed to regularise the model and select only the raw ESG indicators that impact the companies’ default risk. A complete description of the approach is provided in Appendix A.1.

The four specifications are compared in terms of internal validity, assessed based on the model-adjusted R2, and external validity, based on the out-of-sample (OOS) performance. The OOS exercise is conducted with the leave-one-out (LOO) approach, that is, estimating the full model excluding observation i from the dataset and then using the estimates obtained with the $N - i$ observation to predict the i^{th} value. The procedure is explained in detail in subsection 4.1.

2.2 Data

Figure 1 illustrates the correlation between the risk measure in this study, *DtD*, and the aggregate ESG score of the sample companies. It clearly shows that there is no observable correlation between the two and indicates that, at a glance, ESG score is an inappropriate measure of risk.

Figures 2 and 3 illustrate the distribution of companies in the final estimation sample across countries and industries, respectively. On the geographical side, to prevail is the concentration of companies with headquarters located in Germany (155) and the United Kingdom (127), followed by Sweden (88), Switzerland (69), France (62), and Finland (55). On the industrial side, the industrial sector prevails (218), followed by Consumer Discretionary (129) and Information Technology (90).

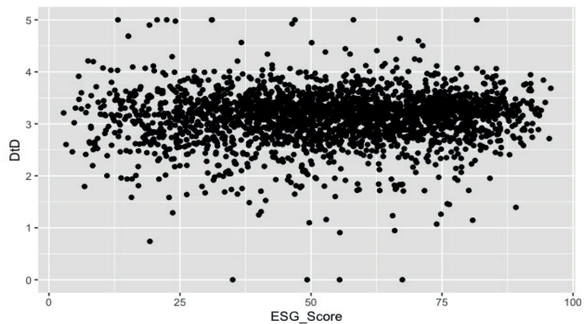

Figure 1: Correlation plot of distance-to-default (DtD) vs aggregate ESG score, EU listed companies, 2021, N=2381.

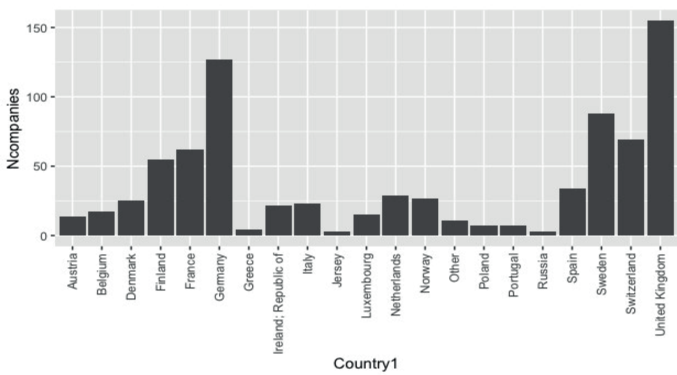

Figure 2: Number of companies by country of headquarters.

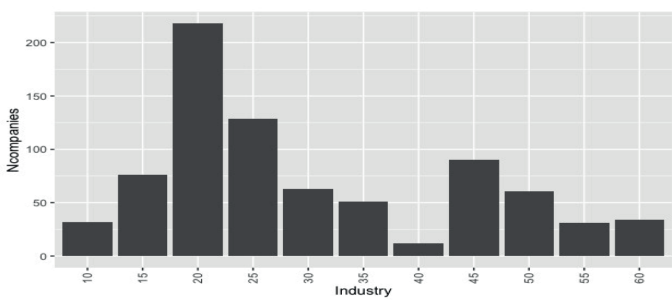

Figure 3: Number of companies by activity industry.

Table 1 provides a descriptive overview of the variables used in the estimation. *DtD* goes from a minimum of 0, approximately close to the default point, to a maximum of 5, indicating a low level of risk. ESG scores show the typical normalised scoring characteristics with a minimum of 0 and a maximum of 100. Only in the case of ESG controversies score, there are companies with the maximum grade. In all the other grades, the maximum is lower than 100, likely because pillar scores are obtained by averaging category scores, and no company graded the maximum in all categories.

Variable	Mean	Min	Max
DtD	3.17	0	5
Size	22.02	16.42	26.73
ROA	0.061	-0.35	1.90
Leverage	0.59	0.07	1.44
Liquidity	1.62	0.12	10.77
TobinQ	1.29	0.034	41.39
Age	50.00	1.00	200.00
ESG Score	64.38	10.66	95.75
ESG Controversies Score	87.97	0.66	100.00
Social Pillar Score	67.00	3.67	97.71
Governance Pillar Score	65.34	10.58	97.73
Environmental Pillar Score	58.61	0.00	99.16

Table 1: Summary statistics. Sample: European listed companies, 2021. Source: Refinitiv.

3. Results

Table 2 summarises the estimation results of equation 1, including the overall ESG score (1), E-S-G pillars sub scores alongside the controversies score (2) and all the category scores, keeping the controversies score (3). In all models, the score coefficient, when significant, has the expected positive sign indicating that a better ESG score is associated with a lower risk (farther from default). However, its size is modest: all else being equal, the expected DtD difference between the best (ESG score = 100) and the worst (ESG score = 0) company is 0.4. In addition, evidence suggests that the G score has no relationship with the DtD, and the same holds for the Controversies score.

Including pillar-specific scores does not substantially improve the explained variance, and the coefficient size is equally small. The model without ESG scores (not shown) explains 37% of the observed *DtD* variability across companies. Including the ESG score among the predictors increases the variance explained to 38% - 39%.

	Dep Var: DtD		
	(1)	(2)	(3)
ESG	0.004***		
	(0.001)		
ESG Controversies Score		0.001	0.001
		(0.001)	(0.001)
Social Pillar Score		0.002*	
		(0.001)	
Governance Pillar Score		0.001	
		(0.001)	
Environmental Pillar Score		0.002**	
		(0.001)	
Resource Use Score			0.002**
			(0.001)
Emissions Score			−0.001
			(0.001)
Environmental Innovation _Score			0.001*
			(0.0004)
Workforce Score			0.0003
			(0.001)
Human Rights Score			−0.001
			(0.001)
Community Score			0.002***
			(0.001)
Product Responsibility Score			0.0005
			(0.001)
Management Score			0.001
			(0.001)
Shareholders Score			−0.001
			(0.001)
CSR Strategy Score			−0.0004
			(0.001)
Adjusted R2	0.379	0.382	0.388

Note: *p<0.1; **p<0.05; ***p<0.01; N= 797

Table 2: Results of the OLS model of Distance-to-Default including ESG information. All models include country and industry dummy variables and the controls in Appendix A.2.

Next to the OLS regression, we consider the LASSO regression of DtD against country and industry dummy, controls, and all the ESG raw scores. A set of candidate models is estimated, one for each possible value of λ (See Equation A.2 in Appendix A.1). Figure 4 shows that all the coefficients are estimated equal to zero for a high value of log(λ). As the value of λ shrinks, a non-zero value is estimated for some of the coefficients. The number of nonzero coefficients becomes larger as the log(λ) values approach their minimum.

The value of λ that minimises the AIC is 0.0138, corresponding to a log(λ) = -4.28 in Figure 4. At this value of λ, 47 coefficients are estimated, which are reported in Table 3. The most critical ESG factor affecting DtD is the "Insider Dealings Controversies Score", which indicates whether the company is under the media spotlight because of a controversy linked to insider dealings and other share price manipulations. The negative coefficient suggests that a company scoring positively on this indicator is closer to the default than a company that is not, all else being equal. This indicator, like many others in the list, suggests that the G pillar is somehow correlated with default, contrasting the result that emerged in Table 2. However, it is important to note that LASSO and OLS coefficients are not directly comparable because of the regularisation bias (Belloni et al., 2013).

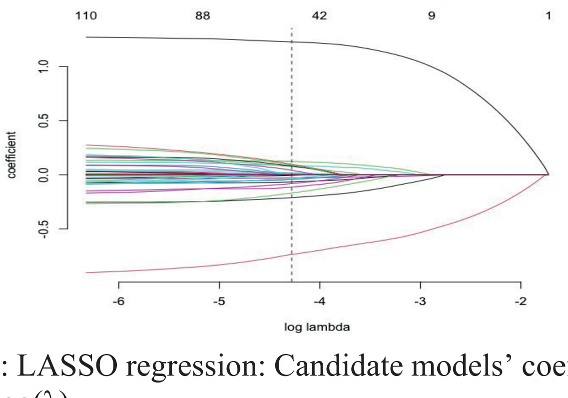

Figure 4: LASSO regression: Candidate models' coefficients against log(λ).

Three primary considerations can be made based on the list of variables in Table 3. The first is that material factors, like those related to GHG emissions or clean electricity or water use, are absent. These variables are characterised by very low reporting percentages and excluded from the estimation database to avoid an excessive sample size reduction. The second is that environmental factors are less relevant than one may think. Only 7 out of the 24 selected factors explicitly refer to the environment. Only one refers to climate change, and they are not among the most important based on the coefficient values. The third and most relevant is that not all factors affect *DtD* as expected. For instance, the ”Environmental Products Score” results suggest that companies reporting on products or services specifically designed to have positive environmental effects are, on average and all else being equal, less risky than companies that do not. In contrast, the ”Climate Change Commercial Risks Opportunities Score” result suggests that companies that report investments in developing new products/services to overcome the threats of climate change are more risky. While the first result is quite intuitive, the second is not. This evidence suggests that the ESG factors weights assigned by score providers may not be consistent, in terms of both expected impact and magnitude, with the purpose of the investor’s ESG integration strategy.

4. Validation

The LASSO regression OOS performance is compared here with those of linear models. Based on the adjusted R2, internal validity - the model’s capacity to predict sample date - improves substantially (from 0.38-0.39 in Table 2 to 0.44), but internal validity is not what ESG practitioners may want to look at. External validity - the model’s capacity to predict new data - reflects the power of new ESG information to predict changes in companies’ riskiness and the possibility to use ESG information about companies for which the DtD is not observed (i.e. non-listed companies) to predict their riskiness. In the first part of the validation exercise, the OOS performance of the different models is compared based on the deviance computed from the LOO estimation. In the second part, different portfolios are constructed screening companies based on the ESG score and the LASSO-based ESGRisk indicator and their financial performances are compared.

Variable	coef
ROA	1,22715
Leverage	-0,73689
Insider Dealings Controversies Score	-0,02941
Executive Compensation Controversies Score	0,00671
Size	0,00364
Improvement Tools Business Ethics Score	0,00131
Responsible Marketing Controversies Score	-0,00094
Voting Cap Percentage Score	-0,00086
Policy Diversity and Opportunity Score	-0,00066
Environmental Products Score	0,00063
Board Structure Policy Score	-0,00058
Policy Business Ethics Score	0,00052
Product Quality Controversies Score	-0,00049
Bribery, Corruption and Fraud Controversies Score	0,00046
Executive Individual Compensation Score	-0,00039
Quality Mgt Systems Score	0,00037
Board Size More Ten Less Eight Score	0,00035
Environment Management Team Score	0,00031
Employees Health and Safety Team Score	0,00030
Environmental Materials Sourcing Score	0,00026
Climate Change Commercial Risks Opportunities Score	-0,00022
Audit Committee Mgt Independence Score	0,00021
Age	0,00019
Sustainability Compensation Incentives Score	0,00015
Environmental Partnerships Score	0,00014
Resource Reduction Targets Score	0,00014
Policy Water Efficiency Score	0,00006
Succession Plan Score	0,00001

Table 3: Indicators selected with LASSO regression and relative coefficients.

4.1. Out-of-sample validation

Consider the general compact form $DtD = \delta' W + \varepsilon$ for the four models in Table 2 and the model estimated with only the LASSO-selected coefficients in Table 3. For each specification, the model is estimated N times, each time excluding observation $i = 1,2,\ldots,N$. The estimated model parameters $\hat{\delta}$ are employed to compute the expected DtD for company i conditional on the company information in the model's right-hand side (W) and the model parameters estimated using all the other $N-i$ companies (Equation 2).

$$DtD_{iLOO} = E(DtD_i | RHS_i, \hat{\delta}_{N-i}) \quad (2)$$

Next, the LOO prediction deviance $\sum(DtD_{LOO} - DtD_i)^2$ is computed. Table 4 reports the estimated deviance for the full sample and the DtD quantiles for the four different models, including:

- the aggregate ESG score (esg);
- the E, S, and G pillars scores (e-s-g);
- category scores (esg-cat);
- raw ESG indicators selected via LASSO (LASSO).

Considering all the observations, LASSO deviance is the lowest, indicating that raw data can more precisely predict a company's expected riskiness than the aggregate ESG score.

Quantile analysis suggests additional insights. The largest deviances are observed in the first and last quantiles, the most and least risky companies, respectively. In these quantiles, in particular, the LASSO deviance is lower, suggesting that using raw ESG data instead of the aggregated ESG score can be particularly effective when the portfolio strategy objective is to identify the least risky companies (positive screening) or avoid investments in the riskiest companies (exclusionary screening).

	esg	e-s-g	esg-cat	LASSO
All obs	113.3525	113.1065	113.2109	106.2962
Q1	53.8	53.7	53.0	49.7
Q2	10.3	10.4	10.5	11.1
Q3	8.02	8.14	8.56	8.19
Q4	41.3	40.9	41.1	37.4

Table 4: OOS Deviance, full sample and DtD quartiles.

4.2. Sustainable portfolios

We finally compare the financial performance of different portfolios made of companies that belong to the top quantile based on either the ESG score (Q1ESG) or the ESGRisk indicator (Q1ESGRisk). The latter is the linear predictor of the LASSO model and can be interpreted as the weighted average of ESG raw scores where coefficients serve as weights determining the relative importance of a specific ESG factor, which is estimated from the data rather than being assumed a priori as usual in ESG aggregation.

We considered different financial performance metrics, namely, the total stock return during fiscal year 2022 (TR-22), the year-to-date total return (TR-YTD) computed on Sept 25th 2023, the 2 year-to-date total returns (TR-2YTD) computed on the same date, and the price performance of the stock compared to the sector it belongs (PCT-SEC). Table 5 presents the average financial performance indicators for four groups: companies excluded from the top quantile according to both indicators, companies in the top quantile according to the ESGRisk only, companies in the top quantile according to the ESG score only, and companies in the top quantile according to both indicators.

Q1ESG	Q1LASSO	TR-22	TR-YTD	TR-2YTD	PCT-SEC
0	0	-0.0299	-0.0197	-0.0480	-0.0145
0	1	0.0535	0.0349	0.0758	0.0168
1	0	0.0135	0.0316	0.0337	0.0427
1	1	0.0817	0.0370	0.130	0.0118

Table 5: LASSOscore and ESG score portfolio performances.

The first three indicators measure total stock returns across different periods and confirm that companies ranking high in both scores have superior stock performance. These companies granted an 8.17% higher return than companies in none of the top quantiles (3.7% and 13% if we consider the return since the beginning of the current and last year, respectively). Looking at the differences between the two scores, returns are higher in companies considered top-raking based on the ESGRisk score compared to the ESG score, indicating that the LASSO algorithm better identifies top-performing companies - in addition to detecting those at lower risk. ESG screening allows the selection of companies with the best financial performance only when this performance is relative to the company's sector (PCT-SEC), suggesting that ESG screening may be effective at the sector level.

5. Conclusion

Many financial institutions still consider positive or negative ESG score screening a reliable strategy for ESG integration that helps minimise risk and maximise returns while supporting the sustainability transition. Extant literature has already demonstrated the weaknesses of ESG scores, on which this argument builds. The primary weakness is the aggregation of information that is not necessarily relevant to the portfolio strategy.

This work leverages the potential of disaggregated raw data using LASSO-based linear model regularisation to detect the ESG factors mostly associated with the default distance. These raw factors are shown to be more capable of explaining and predicting the companies' riskiness. The estimates build a tailored ESG risk score that shows superior performance compared to the aggregated ESG score when used to construct portfolios based on the positive and negative screening criteria.

The study highlights the relevance of raw data for sustainable finance strategy. The extent to which the results may vary among score providers, geographies considered, and industries remains an open field for future research in academia and practitioners.

References

Amel-Zadeh, A., Serafeim, G., 2018. Why and how investors use esg information: Evidence from a global survey. Financial analysts journal 74, 87–103.

Atif, M., Ali, S., 2021. Environmental, social and governance disclosure and default risk. Business Strategy and the Environment 30, 3937–3959.

Bax, K., Sahin, O., Czado, C., Paterlini, S., 2023. Esg, risk, and (tail) dependence. International Review of Financial Analysis 87, 102513.

Belloni, A., Chernozhukov, V., Hansen, C., 2013. Inference on Treatment Effects after Selection among High-Dimensional Controls†. The Review of Economic Studies 81, 608–650. URL:
https://doi.org/10.1093/restud/rdt044, doi:10.1093/restud/rdt044, arXiv:https://academic.oup.com/restud/article-pdf/81/2/608/18394034/rdt044.pdf.

Berg, F., Koelbel, J.F., Rigobon, R., 2022. Aggregate confusion: The divergence of esg ratings. Review of Finance 26, 1315–1344.

Billio, M., Costola, M., Hristova, I., Latino, C., Pelizzon, L., 2021. Inside the esg ratings: (dis)agreement and performance. Corporate Social Responsibility and Environmental Management 28, 1426–1445.

Bonacorsi, L., Cerasi, V., Galfrascoli, P., Manera, M., 2022. Esg factors and firms' credit risk. FEEM Working Paper 36.2022.

Cheng, B., Ioannou, I., Serafeim, G., 2014. Corporate social responsibility and access to finance. Strategic management journal 35, 1–23.

Cho, S.Y., Lee, C., Pfeiffer Jr, R.J., 2013. Corporate social responsibility performance and information asymmetry. Journal of Accounting and Public
Policy 32, 71–83.

Coleton, A., Font Brucart, M., Gutierrez, P., Le Tennier, F., Moor, C.,
2020. Sustainable finance: market practices. European Banking Authority Research Paper.

De Lucia, C., Pazienza, P., Bartlett, M., 2020. Does good esg lead to better financial performances by firms? machine learning and logistic regression models of public enterprises in europe. Sustainability 12, 5317.

EBA, 2020. EBA DISCUSSION PAPER ON MANAGEMENT AND SU-
PERVISION OF ESG RISKS FOR CREDIT INSTITUTIONS AND INVESTMENT FIRMS. Technical Report. European Banking Authority EBA/DP/2020/03.

EC-JRC, OECD, 2008. Handbook on constructing composite indicators: methodology and user guide. OECD publishing.

Ehlers, T., Elsenhuber, U., Jegarasasingam, A., Jondeau, E., 2024. Deconstructing esg scores: investing at the category score level. Journal of Asset Management, 1–23.

Erhart, S., 2022. Take it with a pinch of salt—esg rating of stocks and stock indices. International Review of Financial Analysis 83, 102308.

Erickson, I., Bonne, G., Lawson, D., 2012. Starmine structural credit risk model: overview and global performance. Technical Report. Refinitiv.

Guastella, G., Mazzarano, M., Pareglio, S., Spani, R.C., 2022. Do environmental and emission disclosure affect firms' performance? evidence from sectorial micro data. Eurasian Business Review 12, 695–718.

James, G., Witten, D., Hastie, T., Tibshirani, R., James, G., Witten, D.,
Hastie, T., Tibshirani, R., 2021. Linear model selection and regularization. An introduction to statistical learning: with applications in R, 225–288.

James, G., Witten, D., Hastie, T., Tibshirani, R., Taylor, J., 2023. Statistical learning, in: An Introduction to Statistical Learning: with Applications in Python. Springer, pp. 15–67.

Lanza, A., Bernardini, E., Faiella, I., 2020. Mind the gap! machine learning, esg metrics and sustainable investment. Machine learning, ESG metrics and sustainable investment (June 26, 2020). Bank of Italy Occasional Paper.

Luo, X., Bhattacharya, C.B., 2006. Corporate social responsibility, customer satisfaction, and market value. Journal of marketing 70, 1–18.

Margot, V., Geissler, C., de Franco, C., Monnier, B., 2021. Esg investments: filtering versus machine learning approaches. Applied Economics and Finance 8, 1–16.

OECD, 2020. OECD Business and Finance Outlook 2020: Sustainable and Resilient Finance. OECD Publishing.

Schoenmaker, D., Schramade, W., 2018. Principles of sustainable finance. Oxford University Press.

Appendix

Appendix A.1 Lasso regression

Regularisation avoids too complex models by imposing some penalty on the model dimension. This penalisation is part of the estimation strategy. Let the model written in compact form be DtD = $\delta' W + \varepsilon$, where $\delta = [\alpha,\beta,\gamma]$ and W collects all the right-hand-side variables. If the objective of OLS estimation is to find

$$\hat{\delta} = \operatorname{argmin}\left\{ -\frac{2}{\mathrm{N}} \log \mathrm{L}(\delta) \right\} \tag{A.1}$$

with L being the likelihood function, a regularisation strategy involves adding a penalisation $\lambda \sum_k c(\delta_k)$ to the deviance minimisation problem:

$$\hat{\delta} = \operatorname{argmin}\left\{ -\frac{2}{N} \log L(\delta) + \lambda \sum_k c(\delta_k) \right\}. \tag{A.2}$$

Here, $c(\delta)$ represents the cost of including the variable *k* into the model. There is a clear trade-off in adding complexity to the model: on the one hand, a growing number of variables increases the explained deviance; on the other hand, the cost increases. The balance point is that variables without explanatory power only imply a cost and are left out of the model, accordingly. LASSO regression is a specific regularisation case in which the $c(\delta) = |\delta|$ assumption is made. λ is the penalty weight set between a too-high value determining all $\delta = 0$ and a too-low value determining all $\delta \neq 0$. Hence, LASSO regression produces a set of candidate models, one for each possible value of λ, the best model being selected via *AIC*.

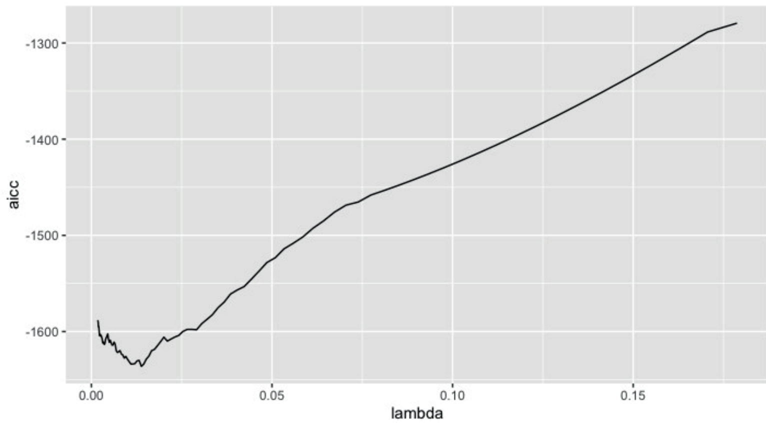

Figure A.5: LASSO regression: plot of model Akaike Information Criterion (AIC) associated with the respective λ value.

Appendix A.2 Sample and control variables

The data used in this study is part of the Refinitiv database. We accessed the database in April 2023 and downloaded financial and non-financial data for all public companies with headquarters in the European continent that recorded a non-zero value of the ESG score in at least one of the last 12 financial years. The original dataset included 2576 companies; the information refers to the last financial year available (2021). For these companies, we accessed information on ESG aggregate scores and raw data in addition to a few structural and financial variables that we used as controls in our study. These include the company size, Returns-on-Asset (ROA), Leverage, Liquidity, Tangibility, Sales, and Tobin's Q, which are acknowledged as primary drivers of risk in the existing literature (Atif and Ali, 2021), in addition to the country of headquarter and the GICS sub-industry. Table A.6 shows the complete list of control variables. Two variables, Tangibility and Sales, show a relatively high number of missing values that would translate into a reduction of the estimation sample size and have been omitted We checked whether the inclusion of these two variables in the model affects the result by comparing the results of the model in equation 1 with and without Tangibility and Sales. The two variables are insignificant ($p = 0.40$ and $p = 0.97$, respectively), while the sign and significance of the other control variables remain unchanged.

The Refinitiv database includes an aggregate ESG score and a pillarspecific score for each E-S-G component. In addition, a separate score concerning ESG controversies is provided. Going at the lower level, each pillar is made by aggregating different category scores for an overall ten categories.

Finally, 213 raw scores cover the ESG universe at the lower disaggregation level. Some of them, however, do not apply to particular industries or, more frequently, only apply to specific industries, leading to an excessive number of NAs in some of the dataset columns. Unfortunately, the dataset does not distinguish the NA cases generated by the sector-specific application of the ESG factor from those generated by its unavailable information. We then consider only the ESG raw factors with a share of NAs lower than 10% in an attempt to include only information available for the majority of companies, and, after that, we take only the complete cases, i.e. the companies for which all the selected factors are observed. The critical point is a clear trade-off between the NAs share threshold and the final number of companies in the estimation sample. Fixing the threshold at 10% leads us to a sample with N =797; lowering it at 5% reduces the number of ESG factors from 99 to 77 but allows N=1053; increasing it at 15% increases the number of ESG factors to 105 but lowers N to 694. It is worth noting that the change in threshold in such a small buffer does not alter the main evidence about the overall relationship between ESG scores and *DtD* and the main conclusion of the work. However, it does affect the LASSO regression results.

Variable	Description
Size	natural logarithm of assets
ROA	ratio of net income over assets.
Leverage	Ratio between accrued debt and assets.
Liquidity	the difference between current and total assets over the difference between current and total liabilities
TobinQ	ratio between the market value of assets and liabilities over the book value of assets and liabilities
Age	Difference between 2023 and foundation year. When the foundation year is absent, Age is replaced with 200.

Table A.6: List of control variables

Printed by
Gi&Gi srl - Triuggio (MB)
September 2024

9788834358696