upna | fec>>

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

school of economics
and business administration
facultad de ciencias
económicas y empresariales
ekonomia eta enpresa
zientzien fakultatea

Facultad de Ciencias Económicas y Empresariales

TRABAJO FIN DE GRADO EN ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESAS

EXPLAINING HIDDEN DATA IN DATABASES: DEFAULT RISK AND DATA AVAILABILITY

Cristina Aragüés Floristán

Pamplona-Iruña,

21 de diciembre de 2022

Módulo Finanzas

Directora: Isabel Abínzano Guillén

Resumen

Vivimos en una época en la que se presiona a las empresas para que divulguen toda la información posible, de forma transparente. El objetivo de este trabajo es examinar si existe una relación entre la cantidad de datos disponibles en bases de datos y su rendimiento financiero, así como el riesgo de impago de la empresa. Además, dado que empresas de todo el mundo han adoptado la elaboración de informes de sostenibilidad, queremos averiguar la relación entre los resultados de factores ESG y la disponibilidad de datos, así como el riesgo de crédito. Utilizamos una muestra de datos de 2.850 empresas estadounidenses desde 2002 hasta 2022, con un total de 57.000 observaciones. Los resultados revelan una relación negativa entre la disponibilidad de datos y el riesgo de impago. Del mismo modo, en lo que respecta a ESG, encontramos una relación negativa entre la puntuación y el riesgo de impago. Esto implicaría que los inversores deberían interpretar una mayor cantidad de divulgación tanto financiera como no financiera como una señal positiva en cuanto a la capacidad de una empresa de pagar sus deudas.

Palabras clave: disponibilidad de datos, rendimiento financiero, riesgo de impago, factores ESG.

Abstract

We live in a time when firms are being pressured to disclose as much information possible, in a transparent manner. The objective of this paper is to examine if there is a relationship between the amount of data availability and the firm’s financial performance and default risk. Furthermore, as sustainability reporting has been worldwide adopted by firms, we want to find out the relationship between ESG factors performance and credit risk. We are using a data sample of 2850 US firms from 2002 until 2022, ending up with 57000 total observations. The findings unveil a negative relationship between data availability and default risk. Similarly, regarding ESG we find a negative relationship between the score and default credit. This would imply that investors should interpret a higher amount of both financial and non-financial disclosure as a positive signal of a firm’s ability to pay back a debt.

Keywords: data availability, financial performance, default risk, ESG factors.

INDEX

1. INTRODUCTION

Transparency can be defined as “the quality of being easy to see through” [1]. When a person is transparent, we think of them as honest, open and trustworthy. Meanwhile, the opposite of transparency is opaqueness, that is, when a person is not see through at all. When it comes to firms, transparency stands for frank, easy to understand financial statements. We live in a time when the economy has become extremely complex, so society expects and demands transparency more than ever. It has become a key component of trust.

This paper focuses on data availability, which is an important part of transparency. Although data availability is a necessary condition for transparency, the latter encompasses more than simple disclosure. For instance, two companies can disclose the same amount of information, but one of them can do so in such a way that investors are not able to properly read the information. Transparency has been defined by Barth and Schipper (2008) as “the extent to which financial reports reveal an entity’s underlying economics in a way that is readily understandable by those using the financial reports”. So, for transparency to be achieved, the financial reports must first give “a true and fair view of the companies”, which is achieved by data availability, and second, they must “be presented in a way that is understandable for their users, which should lead to a decreasing information asymmetry” (Hollanders, 2013). Therefore, while data disclosure basically refers to the “quantity”, transparency concerns both the “quantity” and the “quality” of the information.

Although transparency tends to be associated with positive traits and opaqueness and secrecy with negative ones, truth is finding the balance between the two is a complicated task. We live in a time where intellectual property rights and privacy are in tension with calls for transparency, and in a society that does not like secrecy but is not a fan of full disclosure either. We expect others to give full disclosure while keeping to

[1] Cambridge Dictionary. (2022, September 21). *transparency definition: 1. the characteristic of being easy to see through: 2. a photograph or picture printed on plastic. . .. Learn more.* Retrieved September 24, 2022, from https://dictionary.cambridge.org/dictionary/english/transparency

ourselves as much as we can to protect our competitive advantage, or simply choosing not to reveal what could harm us (Schön, 2006).

In the same way that humans hide their insecurities to appear more valuable to others, firms might intentionally hide information in their financial reports to make sure investors continue to offer up funding. Likewise, just as we post our life highlights on social media and refrain from doing so in our lowest moments, firms might also be likely to report the good news and abstain from disclosing the bad ones. But digging deeper, less information means less certainty for investors: when financial statements are incomplete, investors can never be sure about a company's real assets and true risk (McClure, 2021).

In this paper we study whether firms' information availability is related to their default risk, that is, whether there is a relationship between the amount of information they show in their balance sheets and how risky they are in terms of paying a debt back. To the extent this is true, market participants could interpret a higher information disclosure as a positive signal of a firm's creditworthiness.

We also examine transparency from the point of view of environmental, social, and governance (ESG) criteria, as pressure for their disclosure mounts in the present days. In fact, ESG factors are now more important than ever, and companies are being held accountable for their actions. *"Investors want to invest in businesses that recognize, and are responding to the risks and opportunities ESG practices present, most significantly around climate change"* (PwC Annual Report, 2022). So, ESG reporting is a key component to build trust among stakeholders, as demands are no longer directed towards financial performance but towards other issues that go from diversity to sustainability: reaching net zero carbon emissions, gender inclusion, board diversity, employee benefits, etc.

Additionally, the effects of industry sector, firm size, sector and firm's performance on transparency are examined, in order to see if these characteristics affect their data availability.

The rest of the paper is organized as follows. In the second section we go over the conclusions obtained in previous literature and present our own hypotheses. In the third section we explain our database, present the key variables for the study, which are "data availability", "return on assets", "return on equity", "default risk" and "ESG disclosure score" and go into detail on how we measure them. In the next section we show the results for each of the proposed hypotheses by performing tests of homogeneity of variances and equality of means. The final section gathers the most important conclusions of the paper and exhibits the limitations of the study that give room for future research.

2. LITERATURE REVIEW AND HYPOTHESES

Even though firms are obliged to disclose a considerable amount of financial information, there is still additional data that managers can choose to disclose or not, namely voluntary disclosure (Seran 2022). In a context of information asymmetry where managers hold more information than investors, adverse selection occurs, leading investors to make bad decisions by for instance doing business with riskier firms (Jullobol and Sartmool, 2015).

It is usual that when a firm's performance is not as good as expected, managers use their discretion to conceal it by disclosing less data. This suggests that the less information is available in a company's financial statements, the worse its performance. Or to the contrary, the better a firm's performance, the more information it will show its investors. Consistently, the *"signaling theory"* states that a company will disclose more information than what is required by laws and regulations in order to signal its shareholders that it is healthy (Jullobol and Sartmool, 2015). This is about showing investors that the firm's performance is better than that of other companies (Campbell et al 2001). In this line, Balakrishnan et al. (2014) demonstrate that voluntary disclosure has a positive effect on firm value. Similar results were obtained by a study performed using a dataset of 129 Turkish companies, where voluntary disclosure had a positive effect on firm value (Truong et al, 2022)

Taking this into account, we can suggest that firms with a good financial performance will want to leave this on record on their financial statements, and therefore, the data availability will be higher. This leads us to propose our first hypothesis.

H_1: The better the firm's performance, the more data it will disclose.

In this framework, Melloni et al. (2017) perform a study on a sample of 148 reports of firms that are members of the official IIRC Pilot Programme. They demonstrate that when a firm has a weak financial performance, the company's report will be longer, meaning that it will be less concise, while it will omit information content and bad news at the same time. Their work is supported by Davies and Brennan's view that when a firm's performance is bad, management will make efforts to obfuscate or simply not disclose information, and when it is good, they will be "willing to be forthcoming in their disclosures". This outlook is consistent with the *"signaling theory"*.

Furthermore,_Melloni et al. (2017) use the incomplete revelation hypothesis proposed by Bloomfield (2002) to explain that management has strong incentives to hide bad news, as by making it costly to analyze the reports the market response to the bad news can be reduced. In harmony with this and being "default risk" the event in which a firm can no longer make its payments on its debt obligations, we suggest that firms with a higher default risk will do anything to obfuscate their financial statements, for instance by disclosing less information. Equally, the healthier the firm is in terms of paying back a debt, the more willing it will be to disclose information, both in terms of quantity, which is what concerns this paper, and quality.

Related with this is the "*managerial bad news hoarding hypothesis",* which is defined by Da Silva (2022) as "*the tendency of managers to hide bad news as long as possible from outsiders".*

In harmony with this literature, we propose our second hypothesis:

H_2: The higher the firm's default risk, the less data it will disclose.

In the same line, some authors have been able to prove that increased disclosure has very positive consequences itself. For instance, more disclosure has been shown to be helpful in lowering companies' cost of debt, since part of the cost includes a premium for the uncertainty of the information available about the firm. So, when information asymmetry lowers, uncertainty does too and the cost of capital decreases. In this sense, Leuz and Verrecchia (2000) record a negative relationship between disclosure and the firm's cost of capital (Gow et al 2011). Lambert et al. (2007) also show in a CAPM setting that reduced information uncertainty pushes firms' costs of capital onto the risk-free rate and hence generally lowers costs of capital (Kim et al, 2013). Comparably, Botosan and Plumlee (2002) display that there is a negative relationship between transparency and cost of capital (Gow et al 2011). Given that, we can say that empirical evidence supports the existence of a negative relationship between information disclosure and the cost of capital, as the latter increases when information availability decreases.

On another note, the disclosure of non-financial information is becoming quite relevant, as firms are feeling their stakeholders' pressure to act responsibly in social, environmental and governance areas. The management of the complex relationships firms have with their shareholders and external stakeholders, such as customers, employees, governmental bodies, etc. is explained by the *"stakeholder theory"*, proposed by Freeman in 1984. The theory states that it is no longer about just maximizing shareholders' wealth, as in fact maintaining a good reputation and acting responsibly to meet the different needs of each stakeholder group proves to be very beneficial to the company (Jullobol and Sartmool, 2015). Comparatively, the *"contract theory"* is based on the notion of a social contract between firms and stakeholders, where companies have to embrace responsibilities upon their stakeholders. Given that there is a lot of pressure to do the right thing on ESG issues, and in compliance with the *"signaling theory"* that we mentioned before, those firms with a good performance on ESG criteria are making sure to reflect it on their financial statements to differentiate from others and give a positive signal to external stakeholders.

After a good understanding of the stakeholder and contract theories, we can argue that those firms with the highest amount of financial data available are committed to portraying a realistic image of the firm to its stakeholders, in order to reduce information asymmetry. These companies are likely facing elevated pressures from their stakeholders. And the truth is that if these stakeholders are pressuring the firm to accurately disclose financial information, these stakeholders are not going to let the company get away with murky reporting on the ESG topics, especially after having said that non-financial reporting is being taken into consideration by employers, employees, investors and the wider society. This leads us to develop the following hypothesis:

H_3: The higher the firm's financial data availability, the higher the ESG disclosure

On another note, when the firm's financial performance is good, it can make room for focusing on other topics, such as ESG. The Maslow hierarchy of needs helps understand why firms with a high financial performance might be more likely to have a higher ESG disclosure score: a firm that is focusing on survival will not lose sleep over its self-actualization needs, while a firm that has met its basic needs and is healthy can escalate the pyramid and start worrying about ESG performance.

There are several studies that have examined whether there is a relationship between financial performance and non-financial performance. These studies explain the association through the "*resource-based perspective of firm*", which suggests that firms gain superior performance when they disclose the financial and non-financial resources. For instance, Buallay (2018) obtains evidence from the European banking sector with the aim of finding out if sustainability reporting "ESG" is associated with performance. This study examined 235 banks for ten years (2007-2016) for a total of 2,350 observations, and used ROA, ROE and Tobin's Q as measures of performance. The results of the study show that there is a significant positive impact of ESG on performance. In the study, the author refers to previous literature, such as the research by Steyn (2014), who found that reporting ESG topics contributes to better business with higher financial performance.

For all the above, we are able to propose our fourth hypothesis:

> **H_4**: The weaker the firm's financial performance, the lower its ESG disclosure score will be.

In the same line, firms with a bad financial performance will likely have a higher default risk and we anticipate that they will likely perform worse on ESG than those with a low credit risk. Following the same reasoning as with the previous hypothesis, when a firm has to worry about whether it can pay its debt back or not, its ESG performance will not be the highest priority. Li et al (2022), who examine the implications of ESG practices of Chinese listed firms on their default risk find that firms with a high ESG rating have a low default probability. This allows us to propose our last hypothesis:

> **H_5**: The higher the firm's default risk, the lower its ESG disclosure score will be.

3. DATA AND RESEARCH METHODOLOGY

3.1. Database

Once the most relevant literature has been reviewed and we have clarified some concepts, we can move on to perform our own empirical analysis with the objective of testing the stated hypothesis. But before presenting our results, it seems pertinent to describe our database and present the key variables for the study.

We obtained data of 2,850 US firms from 2002 until 2022 from the Eikon Refinitiv database. The information obtained includes market and accounting information. We impose the existence of at least data for the market value variable, and also extract the information, available or not, of other variables. Specifically, we select the following variables, which are the ones included in the most commonly used measures of default risk, such as the Black-Scholes-Merton measure and Altman's Z-score: short-term debt, long-term debt, EBIT, retained earnings, total assets, net income, working capital, market to book value ratio, current assets, current liabilities, operating income and funds from operations. The maximum value possible is 13 and the minimum is 1. Since we have imposed the existence of the Market Value variable, if the company has no data for that variable, the data availability variable is missing for that firm-year observation.

In addition, we also obtained their Standard Industrial Classification (SIC) code in order to group them and see whether there were differences among industries. The divisions were: Agriculture, Forestry and Fishing; Mining; Construction; Manufacturing; Transportation, Communications, Electric, Gas and Sanitary Services; Wholesale Trade; Retail Trade; Finance, Insurance and Real Estate; Services; and Administration.

3.2. Key variables

As it has been previously mentioned, the principal objective of this paper is to examine the relationship between data availability, financial performance and default risk, with also a focus on ESG. For this, it is appropriate to present the key variables for the study and explain how we will measure them.

3.2.1. Data availability

With the aim of measuring data availability, we have counted the number of variables available for each firm for each year, taking into account the aforementioned thirteen variables. Since we require at least information for Market Value variable, the number of variables available ranks from 1 to 13. In case a company doesn´t have information on Market Value, the variable of data availability is an empty data. In Table 1 we can see this information both, by company and by year.

Table 1. Number of variables by firm-year

By company			By year		
Maximum	Minimum	Mean	Maximum	Minimum	Mean
13	1	10,1679071	13	1	11,2148008

Table 1 shows the number of variables available by firm and by year. The minimum for both is 1, as we established the condition that market value should be available. The maximum is 13, meaning there are companies with data available for all the different variables. On average 10.17 variables were obtainable for each firm, while 11.21 variables were accessible each year.

Furthermore, in Table 2 we show the same information but by variable. We can notice that the least available variable is EBIT, while the most available variable after market value is Net Income.

Table 2. Availability of data by variable by company-year

	By company			By year		
	Maximum	Minimum	Mean	Maximum	Minimum	Mean
Market value	20	1	10.5522807	1676	1361	1503.7
Total Assets	20	0	10.72280702	1618	1380	1528
Short Term Debt	20	0	10.43964912	1586	1355	1487.65
Long Term Debt	20	0	10.67789474	1609	1377	1521.6
EBIT	20	0	6.294385965	1264	58	896.95
Retained Earnings	20	0	9.991578947	1525	1298	1423.8
Net Income	20	0	10.9077193	1647	1412	1554.35
Working Capital	20	0	9.241403509	1393	1215	1316.9
Net Sales	20	0	10.91789474	1648	1412	1555.8
Current Assets	20	0	9.250175439	1398	1215	1318.15
Current Liabilities	20	0	9.251578947	1396	1216	1318.35
Operating Income	20	0	10.88491228	1637	1411	1551.1
Cash from Operations	20	0	10.79368421	1622	1397	1538.1

Table 2 shows the amount of data that is available by company and by year for thirteen different variables. On the left side of the table, we have the amount of data available in years amount for each variable by company.

For instance, regarding 'market value', the minimum amount of data available is one. This means that there are companies that only have data available regarding their market value in one year.

If we look at the "Mean" column, we observe that there are noticeable differences between variables. As an illustration, we observe that EBIT, Earnings Before Interest and Taxes is the variable for which the least amount of data is available by company throughout the years, with an average of "6.29". On the other hand, Net Sales is the variable with the highest amount of data available throughout the years, with an average of "10.92". Why are there more available data for certain variables than for others?

The right side of the table shows the number of data available by year for each variable. For instance, the year with the maximum amount of data available regarding market value had 1676 data, that is, the market value of 1676 companies was available. Once again, when looking at the mean Earnings Before Interest and Taxes is the variable for which the least amount of data was available, with an average of 897 companies' EBIT at hand. Contrastingly, Net Sales was the variable with most available data by year, following the same pattern we observed with the availability of data by companies.

3.2.2. Measure of performance

The two variables that we will use to measure performance are ROA and ROE, standing for "Return on Assets" and "Return on Equity", respectively.

On the one hand, ROA measures the profitability of a firm, in relation to its total assets. It is also called "operational profit", meaning it is the profit the firm obtains from its operations. It is calculated by dividing earnings before interest and taxes by the total assets of the firm. The higher the ratio, the more efficiently the company is using its resources.

On the other hand, ROE measures the profitability of a firm in relation to its equity, that is, its total assets minus debt. It is also called "financial profit". It is calculated by dividing net income by the total assets of the firm. The higher the ratio the more efficient a company is at generating income from its net assets.

3.2.3. Measure of default risk

In order to measure default risk, we use a market-based measure of credit risk, since this kind of measure has been proven more accurate than accounting-based models and credit rating (Hillegeist el al., 2004, Gharghory et al., 2006). In terms of availability, this measure overperforms other measures, such as the Credit Default Swaps spreads (Hilscher and Wilson, 2017; Abinzano et al., 2020).

Specifically, we use the model proposed by Bharath and Shumway (2008), which is a naïve version of the structural model known as Black-Scholes-Merton (BSM) measure to avoid the iterative process needed to implement the BSM measure. This way, Bharath and Shumway (2008) derive the following expression for the probability of default of a company:

$$P_{i,t}^{def} = N\left(-\frac{ln\frac{E_{i,t}+D_{i,t}}{D_{i,t}}+\left(r_{i,t}-\frac{\sigma_{A_{i,t}}^2}{2}\right)(T-t)}{\sigma_{A_{i,t}}\sqrt{T-t}}\right) \tag{1}$$

with:

$$\sigma_{A_{i,t}} = \frac{E_{i,t}}{E_{i,t}+D_{i,t}}\sigma_{E_{i,t}} + \frac{D_{i,t}}{E_{i,t}+D_{i,t}}(0{,}05 + 0{,}25\sigma_{E_{i,t}}) \tag{2}$$

and where $E_{i,t}$ is the market capitalization of the firm *i* at *t*, $D_{i,t}$ is the face value of the debt, $r_{i,t}$ is the past annual return of the firm, $\sigma_{E_{i,t}}$is the annual volatility of the stock value, $\sigma_{A_{i,t}}$is a proxy for the volatility of the market value of the firm's total assets, *T-t* is the time to maturity, and N(·) is the cumulative probability of the Normal distribution (0,1). As in other works (Crouhy et al., 2000; Vassalou and Xing, 2004; Gharghori et al., 2006), we take *T-t = 1* year, and the face value of debt as the sum of short-term debt plus half of long-term debt.

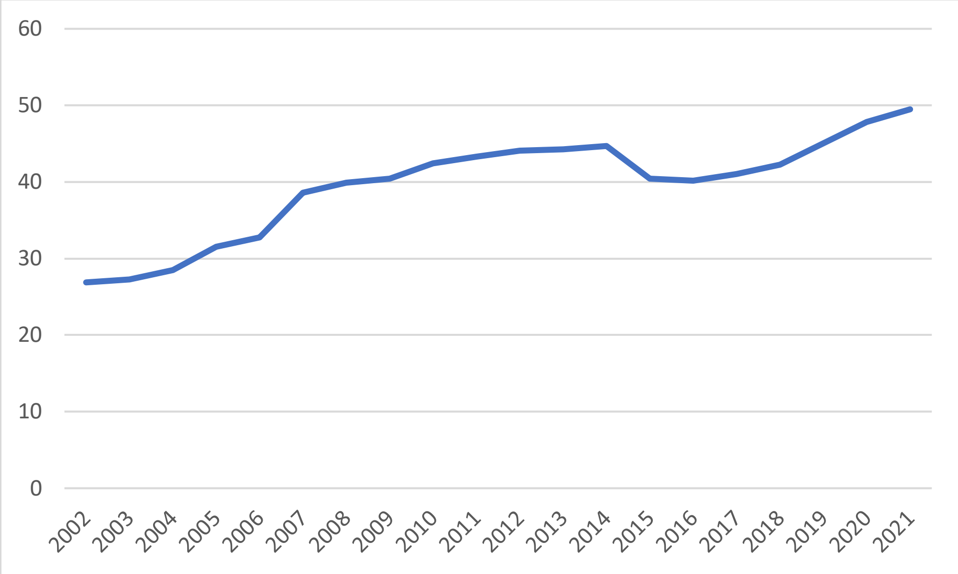
3.2.4. ESG disclosure score

ESG scores are obtained from the Eikon Refinitiv database. They are calculated by considering the companies' performance, commitment, and effectiveness among 10 main themes: emissions, resources, environmental product innovation, workforce, human rights, community, product responsibility, management, shareholders, and CSR strategy. They take into account the company's sector (environmental and social) and country of incorporation (governance) to obtain the performance of ESG factors.

The score can range from 0 to 100, with 0 indicating "poor relative ESG performance and insufficient degree of transparency in reporting material ESG data publicly" and 100 meaning "excellent relative ESG performance and high degree of transparency in reporting material ESG data publicly". Firms with an ESG score close to 0 can be called "ESG laggards", while those with a score close to 100 can be called "ESG leaders".

Graph 1 shows the evolution of the average ESG score from 2002 up to 2021. The trend shows how ESG matters are gaining importance, as the average score has gone from 27 up to 50.

Graph 1. Evolution of the average ESG disclosure score



Although the score is slowly going up, thanks to awareness about ESG topics, its average is still below 50.

Now that we have explained our database and the key variables for the paper, we can move on to the next Section.

4. RESULTS

In the present section we will present the results for each of the formulated hypothesis, with the first part being an analysis of how data disclosure varies depending on the characteristics of the firm.

But before moving on to the analysis of each hypothesis, in the following table we provide descriptive statistics of the sample for the most relevant variables.

Panel A of Table 3 shows the statistics mean, standard deviation, minimum, median and maximum for the relevant variables of our study, including "ROA", "ROE" and "Credit risk". It should be highlighted that the mean for "data availability" is quite high: 11.2163, considering the maximum for this variable is 13.

The first five variables correspond to firm characteristics: size (in euros, measured with market value and Total Assets), leverage, volatility of equity and Market to Book Value ratio. The next variables are the ones we use to examine our hypothesis and the ones we explained in 3.2: data availability, ROA and ROE (to measure performance), credit risk, and ESG disclosure score.

In Panel B we have the correlation matrix. We observe that data availability is positively correlated with market value and volatility of equity, and also with ESG score, while negatively with leverage, credit risk and ROE.

Table 3. Sample descriptives

Panel A: Main descriptives

	Mean	Standard deviation	Minimum	Median	Maximum
Size (MV)	22,85322	23,98028	11,15625	21,33564	26,96137
Size (TA)	22,88345	24,05815	0	18,80772	27,40508
Leverage	0,32392	0,26581	0	0,30096	10,08163
Volatility of equity	0,23813	0,25974	0	0,18009	4,07093
MTBV ratio	4,48154	211,91032	-31556,55	2,27000	11217,68
Data availability	11,21630	3,23484	1	12	13
ROA	0,00019	0,01090	-0,00293	0,00008	1,41375
ROE	0,06434	14,02352	-1520,024	0,05331	1659,37143
Credit risk	0,04330	0,16610	0	1,09E-16	1
ESG score	41,54581	19,82766	0,59	38,63	94,43

Panel B: Correlation Matrix

	Size (MV)	Size (TA)	Leverage	Volatility of equity	MBTV Ratio	Data availability	ROA	ROE	Credit risk
Size (TA)	-0.0338***								
Leverage	0.0093	0.0841***							
Volatility of equity	-0.2954***	-0.0284***	0.0528***						
MBTV Ratio	-0.0132**	-0.0203***	-0.0207***	-0.0091*					
Data availability	0.1816***	-0.1463***	-0.1774***	0.0433***	-0.0081				
ROA	-0.0018	-0.0354***	-0.0136*	0.0016	0.0152**	-0.0049			
ROE	-0.0041	-0.0168***	-0.0223***	0.001	0.0351***	-0.0209***	-0.0991***		
Credit risk	-0.2646***	0.03***	0.2626***	0.3777***	-0.0176***	-0.0731***	-0.1058***	-0.0023	
ESG	0.5039***	-0.1801***	0.0086	-0.1634***	0.0108	0.138***	0.0743***	0.0143*	-0.0631***

4.1. Characteristics of the companies depending on data ability

How are the firm's characteristics related to amount of data it discloses? Do bigger firms disclose more or less data than smaller firms? Does their leverage ratio also have an effect on this?

In this section we analyze whether "Size", "Leverage", "Volatility of equity" and "MTBV ratio" have an effect on the amount of data the firm discloses. This has been achieved by examining if there are statistically significant differences in the amount of disclosure among companies with different characteristics.

Table 4 shows the average values for our most important variables when availability is high, that is, when the number of variables is higher than 7, and the average values when availability is low, when the number of variables equals 7 or lower.

Table 4. Company characteristics high-low availability

	High availability	Low availability
Size (MV)	21,43735748	20,86074653
Size (TA)	18,70519913	16,13580787
Leverage	0,319545844	0,049593466
Volatility E	0,332102057	0,290739394
MTBV ratio	3,445394742	6,43484082

This table shows the firm's characteristics in relation to its data availability. Since the minimum amount of variables the firm can disclose is 1 (Market Value) and the highest is 13, we split it at 7 (the median of these values) and took this value as the threshold to determine which firms have high or low availability. On the left we see the average values for the companies that disclose information for more than 7 variables. On the right, we see the average values for the companies that disclose information for 7 or less variables.

We can already see from the results of the table that there are differences in the average values of the companies with high availability and low availability.

The first thing we checked was whether it was reasonable to assume that the companies with high availability have the same population variance as those with low availability. In order to test this, we created a Dummy variable for Data Availability, where it equals 1 when the variable is higher than 7 (the median of the maximum and the minimum, since the number of variables ranks from 1 to 13) and 0 when it equals 7 or less. Considering equal or different population variances, we later tested the equality of means to see whether the firms' characteristics have a significant effect on disclosure.

Table 5. Company characteristics variance and mean test for high-low availability

	Statistic	High availability	Low availability	Difference
Size (euros)	Variance	2,79108	2,756410	0,03467
	Mean	21,43740	20,860700	0,57670***
Size (TA)	Variance	6,5137	5,41501	1,0987
	Mean	18,7052	16,13580	2,5694***
Leverage	Variance	0,05807	0,01813	0,03994***
	Mean	0,31955	0,049594	0,269953***
Volatility E	Variance	0,065640	0,06199	0,0037**
	Mean	0,33210	0,290739	0,04136***
MTBV ratio	Variance	60628	5939,97	54688,03***
	Mean	3,44539	6,43484	-2,98945*

The table shows the tests for equal variances and means for the variables size (market value and total assets), leverage, volatility of equity and MBTV ratio when data availability is high or low . Data availability is the independent variable, and size, leverage, volatility of equity and MTBV ratio are the dependent variables. ***, **, and * indicates statistical significance at 1 percent, 5 percent, and 10 percent levels. High availability (dummy) takes value 1 if company discloses information for more than 7 variables and 0 otherwise.

The results indicate that we can only assume equality of population variances for firm size (measured with the total assets variable). For the rest of the variables we cannot, with a significance of 1% for all of them.

Regarding equality of means, we see that there is a positive, statistically significant relationship between both, size and leverage with data availability, both at the level of 1%. The relationship between the completeness of disclosure and volatility of equity is also positive and statistically significant, but at the level of 5%. In relation to the market to book ratio, the relationship is negative and still statistically significant, but this time at the level of 10%.

In the following paragraphs we will go into more detail with every characteristic and try to explain the why to these relationships.

Beginning with size, this variable is related to the structure of the company. It can be calculated by computing the natural logarithm of market value or of total assets. Regarding its relationship with data availability, we see size of companies with more data availability is statistically higher at a 1% level. There are various reasons why large companies might be disclosing more information. These reasons are analyzed on a study performed on Swedish listed companies by Andersson and Folkare (2015) which finds a positive relationship between firm size and disclosure. They think it is logical that bigger firms disclose more information, as "*large companies have the resources to disclose more and are also more scrutinized by the public, since they have a larger impact on society.*" The impact of firm size confirms the theoretical assumptions made from the "stakeholder theory". In their study, Andersson and Folkare (2015) collect the conclusions of several previous studies that have found firm size to have a positive effect on disclosure, such as Watson et al. (2002), Guthrie et al. (2006), White et al. (2007) and Broberg et al. (2010). For all this, we can conclude that the relationship between firm size and voluntary disclosure is positive and significant. Upon analyzing our database in detail, we see that many of the firms that are disclosing the maximum amount of variables, 13, are some of the largest. Some examples are "Johnson & Johnson", "Procter & Gamble", "Walmart", "McDonald's", "Mastercard", "Cocacola" and "Pfizer". These firms have many stakeholders to satisfy and big pressures to disclose their information transparently.

Leverage is also a structure-related variable. It is calculated by dividing total debt by total assets and it shows how much the company is financed by debt. We see that the effect of leverage on data disclosure is statistically significant at a 1% level. Results show that companies with a higher leverage show a more complete disclosure than those with a lower leverage. Our findings are consistent with the study performed by Seran (2022) on 434 Indonesian listed companies, according to whom companies should disclose more detailed information when they have a higher dependence on creditors' so they can meet their needs.

Volatility of equity has been calculated as the standard deviation of the previous four years' annual return of the company. It measures the dispersion of returns. Higher volatility usually implies a higher risk. We find that the relationship between the volatility of equity and data disclosure is positive and statistically significant at a 1% level, with higher volatility for companies with higher availability

Finally, the book-to-market ratio is calculated by dividing the company's stocks' price in the market by its book value. Results show that firms with a lower amount of information show a higher ratio, with a level of significance of 10%. So, there is a negative but weakly significant relationship between data availability and the book-to-market ratio.

For all the above, we can conclude that the size of the company, its leverage, volatility of equity and MTBV ratio all have a significant relationship with data availability.

In addition, we gathered in a table the results for the average amount of variables available in a company's statements depending on industry type, as determined by their Standard Industrial Classification (SIC Code).

Table 6. Number of variables available by industry type

Industrial sector	Average # variables	n
Agriculture, forestry & fishing	11,67857143	6
Mining	11,37453581	189
Construction	12,12643678	50
Manufacturing	11,91676532	902
Transportation, Communications, Electric, Gas, And Sanitary Services	11,64826226	347
Wholesale Trade	11,73542117	87
Retail Trade	11,94048205	174
Finance, Insurance and Real State	9,174356596	454
Services	11,57433319	531
National security	12,875	1
		2741

The table shows the average number of variables the firms belonging to each of the sectors are disclosing, taking into account the rank can go from 1 to 13. The right column shows the number of firms belonging to each sector. Even though our data base consists of 2850 firms, we only have data regarding the SIC code of 2741.

We observe that the firms in the Financial sector, consisting of 454 companies, are less likely to disclose information, while firms in the National security, although the sample size is only one, and the Construction sector, consisting of 50 firms, are the most likely to do so. Within the Financial sector, as shown in table 7, we observe that Depository institutions (for instance banks), consisting of 7 entities, are the most likely to disclose information with a mean of 10.65 variables available, while the Holding offices, consisting of 312 companies, and the Non-depository credit institutions (for example insurance companies), consisting of 11 companies, are the least likely to do so, with the respective means of 8.16 and 8.82 variables available.

Table 7. Number of variables available in the Financial Sector

Finance, insurance and Real State	Average # of variables	n
Depository institutions	10,65517241	7
Non-depository credit institutions	8,821428571	11
Security And Commodity Brokers, Dealers, Exchanges, And Services	8,666666667	3
Insurance carriers	9,328413284	22

Insurance Agents, Brokers, And Service	9,117647059	3
Real Estate	9,373655914	96
Holding and other invest. Offices	8,157232704	312
		454

Now that we have seen the relationships between company characteristics and sector and the amount of financial information they disclose, we can move onto the next section and begin going over each of our proposed hypotheses.

4.2. Performance of the companies depending on data availability

The present section concerns analyzing whether our first raised hypothesis H_1: "*The better the firm's performance, the more data it will disclose*" holds or not. In the same way, we did to check if firm characteristics were related data disclosure, we performed tests of homogeneity of variances to see if we could assume equal population variances and later on carried out tests of equality of means.

Table 8. ROA and ROE variance and mean test for high-low availability

	Statistic	High availability	Low availability	Difference
ROA	Variance	Entry not valid	Entry not valid	
	Mean	Entry not valid	Entry not valid	
ROE	Variance	95,865	1687,8700	-1592,0051***
	Mean	0,015497	12,9887	-12,9732***

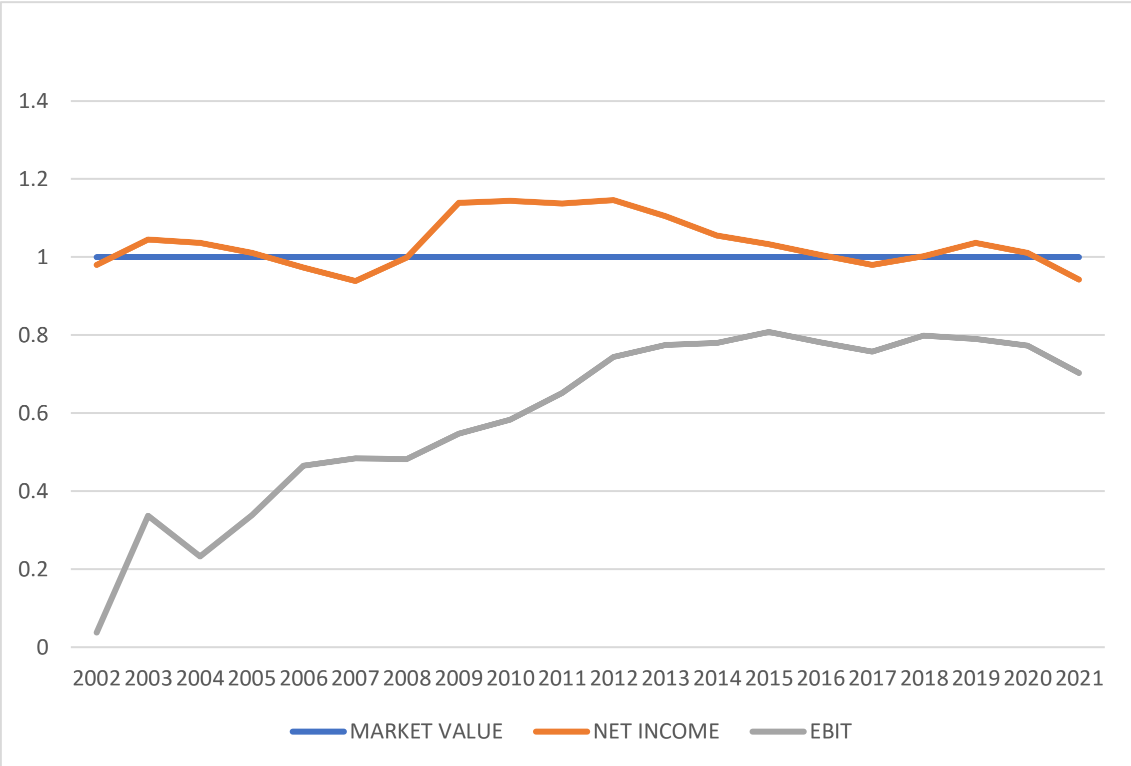
The table shows the tests for equal variances and means for the variables ROA and ROE when data availability is high or low. Data availability is the independent variable, and ROA and ROE are the dependent ones. ***, **, and * indicates statistical significance at 1 percent, 5 percent, and 10 percent levels. High availability (dummy) takes value 1 if company discloses information for more than 7 variables and 0 otherwise.

Regarding ROA, we were not able to perform the test for equal population variances and means since the sample size for low availability was only 1. This means that among all the firms that present a low data availability, we only had data regarding the return on assets of one firm, in one year.

In relation with ROE, equality of population variances cannot be assumed. Regarding equality of means, we observed that a higher ratio is related with lower disclosure. More specifically, firms with a high data availability show an average ROE of 0.015497, while firms with a low data availability show an average ROE of 12.9887. These results are statistically significant at the level of 1%.

Since we use EBIT to calculate ROA and Net Income to calculate ROE, we noticed the reason why we were not able to test for homogeneity of variances and means for ROA on Gretl is that the number of total observations we have for the EBIT quite low, especially compared to Net Income. This is something we observed when we explained Table 2, as EBIT was the variable for which the least amount of data was available. We only had 17939 total observations for EBIT, out of 57000 (highest amount possible, as we have a sample size of 2850 firms during 20 years). Graph 2 shows the percentage of observations of EBIT and Net Income per year compared to the number of Market Value observations.

Graph 2. Percentage of EBIT and Net Income observations



For some reason, companies are showing less data for Earnings before Interests and Taxes than they do for Net Income, especially from 2002 until 2012. From 2015 until 2021, there is about twenty percent less data for EBIT than Net Income, but the increases and decreases in availability follow the same trend through the years. EBIT is relevant because it nulls the effects of the different capital structures and tax rates used by different companies. By excluding both taxes and interest expenses, EBIT is an honest way to show the company's ability to make profit and makes it easier to compare among firms (Adiloğlu and Bengu, 2017). These authors argue that EBIT is a more important measure of the firm's performance than Net Income. This is because when Net Income is calculated, companies decide on non-operating matters such as interest expenses and tax rates, which are management and governmental decisions that have little to do with the firm's real performance, which is the firm's ability to generate profit. Our results show that although EBIT disclosure has increased through the years, some companies are still not declaring their real performance and are only showing Net Income as a measure of performance. So, although the companies' attitude is promising, they still have a long way to put their EBIT disclosure level on the same level as Net Income is.

It is also worth noting that firms' disclosure practices vary through time as do the economic conditions. For example, we can see in the graph that in 2020 both the Net Income and EBIT disclosure decreases, possibly as a result of a weaker financial performance associated with Covid-19.

Because we were not able to perform our tests for ROA with the dummy variable we created for availability, we did a complementary analysis by setting a higher threshold so that High Availability (dummy) takes value 1 if company discloses information for more than 10 variables and 0 otherwise. By setting the threshold on 10 instead of 7, we expect to have more observations when the dummy takes value 0 (Low availability) and therefore be able to extract more robust conclusions for our first hypothesis. Once the new dummy variable is created, we run the additional tests to obtain better conclusions.

Table 9. ROA and ROE variance and mean test for high-low availability

	Statistic	High availability	Low availability	Difference
ROA	Variance	0,000132386	0,000000001	0,000132385***
	Mean	0,000202875	0,000036209	0,000166666**
ROE	Variance	109,510000000	36,039000000	73,471000000***
	Mean	0,006751310	0,307925000	-0,301173690***

The table shows the tests for equal variances and means for the variables ROA and ROE when data availability is high or low. Data availability is the independent variable, and ROA and ROE are the dependent variables. ***, **, and * indicates statistical significance at 1 percent, 5 percent, and 10 percent levels. High availability (dummy) takes value 1 if company discloses information for more than 10 variables and 0 otherwise.

First, we saw that homogeneity of population variances could not be assumed for neither ROA nor ROE, both with a level of significance of 1%. As for equality of means, we found that Return on Assets for those firms with a low availability is lower than for those with high availability, 0.000036 compared to 0.00020, with a level of significance of 5%. Contrastingly, the Return on Equity for companies with a low data availability is much higher than for the firms with a high availability, 0.31 compared to 0.007, with a level of significance of 1%. The previous test also showed a negative relationship between the two variables, but now the difference between means is even higher.

This shows that there is a positive relationship between availability and ROA, but a negative one between availability and ROE.

In conclusion, we cannot accept nor reject H_1, as operational performance (ROA) is positively associated with data availability, but financial performance (ROE) is negatively associated with it.

4.3. Default risk of the companies depending on data availability

This section concerns the analysis of our second hypothesis, that is, whether a firm's default risk is related with the amount of information it discloses in its financial statements, $H_{2:}$ The higher the firm's default risk, the less data it will disclose.

Table 10. Credit risk variance and mean test for high-low availability

	Statistic	High availability	Low availability	Difference
Credit risk	Variance	0,0276923	0,0894686	-0,0617763***
	Mean	0,0434715	0,155386	-0,1119145

The table shows the tests for equal variances and means for the Credit Risk variable when data availability is high or low. Data availability is the independent variable, and Credit Risk is the dependent variable. ***, **, and * indicates statistical significance at 1 percent, 5 percent, and 10 percent levels. High availability (dummy) takes value 1 if company discloses information for more than 7 variables and 0 otherwise.

We start off the base that equal population variances cannot be assumed for default risk. Moving onto our test for equal population means, we obtain that our results are not statistically significant. Therefore, based on this test cannot say that default risk has a significant effect on data availability.

In the same way we performed a complementary analysis by increasing the data availability threshold from 7 to 10 to examine the relationship between performance and data availability, we did that same analysis with credit risk in order to check whether we could determine a significant relationship between credit risk and data disclosure.

Table 11. Credit risk variance and mean test for high-low availability

	Statistic	High availability	Low availability	Difference
Credit risk	Variance	0,0254149	0,0426231	-0,0172082***
	Mean	0,0404073	0,0642817	-0,0238744***

The table shows the tests for equal variances and means for the Credit Risk variable when data availability is high or low. Data availability is the independent variable, and Credit Risk is the dependent variable ***, **, and * indicates statistical significance at 1 percent, 5 percent, and 10 percent levels. High availability (dummy) takes value 1 if company discloses information for more than 10 variables and 0 otherwise.

We found that homogeneity of variances could not be assumed, with a level of significance of 1%. Regarding equality of means, we found credit risk having a negative effect on the disclosure of data. We should highlight that when our threshold for creating the Dummy variable for availability was 7, the mean for the firms with low data availability goes from 0,155386 to 0,0642817. The most relevant contribution in relation with our previous analysis is that these results are statistically highly significant, at the level of 1%.

For all the above, the results of our test support H2, which claims that the higher the firm's default risk, the lower data it will disclose. The results are in line with the incomplete revelation hypothesis proposed by Bloomfield (2002).

4.4. Relationship between data availability and ESG disclosure score

Now we can move on to presenting the results for the proposed hypothesis involving ESG.

Beginning with H_3, that is, "The higher the firm's financial data availability, the higher the ESG disclosure will be", it is based on the belief that the pressures endured by the firms with high data availability to disclose financial information must be similar to the pressures they bear regarding non-financial information.

To test this hypothesis, we followed the same method as with H_1 and H_2: we tested for homogeneity of variances and equality of means.

Table 12. ESG score variance and mean test for high-low availability

	Statistic	High availability	Low availability	Difference
ESG score	Variance	393,2610	179,05800	214,2030***
	Mean	41,6894	26,75370	14,9357***

The table shows the tests for equal variances and means for the ESG disclosure score variable when data availability is high or low. Data availability is the independent variable, and ESG score is the dependent variable ***, **, and * indicates statistical significance at 1 percent, 5 percent, and 10 percent levels. High availability (dummy) takes value 1 if company discloses information for more than 7 variables and 0 otherwise.

We reject the null hypothesis that population variances are equal, and taking these results into account, we test for equality of means. We found that there is a significant difference between the ESG disclosure score of those firms with a high data availability and those with a low availability. More specifically, firms with a high financial information disclosure show an average ESG core of 41.6894, as opposed to the average score of 26.7537 of firms with a low financial information availability. As a consequence, we cannot reject Hypothesis number 3, confirming that the amount of financial information disclosed has a positive and significant relationship with the ESG score.

4.5. ESG disclosure score of the companies depending on financial performance

This section concerns the analysis of H_4, which claims that "The weaker the firm's financial performance, the lower its ESG disclosure score will be."

In the same way we created a dummy variable for data availability to examine Hypothesis 1, 2 and 3, and we created dummy variables for ROA and ROE. The way we did this was by splitting it at the median value for each year so that the values higher than the median for each year take the value 1 and those equal or lower take the value 0.

Table 13. ESG score variance and mean test for high-low ROA

	Statistic	High ROA	Low ROA	Difference
ESG score	Variance	394,46400	275,631000	118,83300***
	Mean	44,30470	34,594600	9,71010***

The table shows the tests for equal variances and means for the ESG score variable when ROA is high or low. ROA is the independent variable, and ESG score is the dependent variable ***, **, and * indicates statistical significance at 1 percent, 5 percent, and 10 percent levels. High ROA (dummy) takes value 1 if company has a ROA higher than the median that year and 0 otherwise.

Table 14. ESG score variance and mean test for high-low ROE

	Statistic	High ROE	Low ROE	Difference
ESG score	Variance	400,4720	331,81500	68,6570***
	Mean	42,6933	36,75070	5,9426***

The table shows the tests for equal variances and means for the ESG score variable when ROE is high or low. ROE is the independent variable, and ESG score is the dependent variable ***, **, and * indicates statistical significance at 1 percent, 5 percent, and 10 percent levels. High ROA (dummy) takes value 1 if company has a ROE higher than the median that year and 0 otherwise.

To test this hypothesis, we followed the same method as with the previous three. We first performed tests of homogeneity of variances to see if we could assume equal population variances. Population variances could not be assumed to be equal for neither ROA nor ROE. In analyzing the difference of means, the results supported a statistically significant relationship, at a level of 1%, between high performance and a higher ESG disclosure score, therefore supporting H_4, independently on the measure of the performance.

More specifically, we see that firms with a high Return on Assets have an average ESG disclosure score of 44.3047, in contrast with the average score of 34.5946 of the firms with a low ROA. In the same manner, we find that firms with a high Return on

Equity have an average ESG disclosure score of 42.6933, in contrast with the mean score of 36.7507 of the firms with a low ROE.

Our results are in line with the studies performed by Buallay (2018) and Steyn (2014), who found a positive association between ESG reporting and financial performance.

4.6. ESG disclosure score of the companies depending on credit risk

This section analyzes our last proposed hypothesis H_5: "The higher the firm's default risk, the lower its ESG disclosure score will be."

To test the hypothesis, we created a Dummy variable for credit risk. We did this by splitting it at the median value for each year so that the values higher than the median credit risk for each year take the value 1 and those equal or lower take the value 0.

Table 15. ESG score variance and mean test for high-low Credit Risk

	Statistic	High credit risk	Low credit risk	Difference
ESG score	Variance	393,2840	441,45800	-48,1740**
	Mean	42,6728	45,09760	-2,4248***

The table shows the tests for equal variances and means for the ESG score variable when the firm's credit risk is high or low. Credit risk is the independent variable, and ESG score is the dependent variable ***, **, and * indicates statistical significance at 1 percent, 5 percent, and 10 percent levels. High Credit risk (dummy) takes value 1 if company has a default risk higher than the median that year and 0 otherwise.

Equal population variances could not be assumed, at a level of significance of 5%. As for equality of means, we reject the null hypothesis that they are equal, as results show a statistically significant difference at the level of 1%. We observe that firms with a low credit risk have a higher average ESG disclosure score than firms with a high risk. More specifically the former have an average score of 45.0976, while the latter show a mean score of 42.6728.

In conclusion, we cannot reject H5. Non-financial performance is negatively associated with credit risk, meaning that those firms with a lower credit risk are performing better on ESG factors.

4.7. Implications

After examining the results of all the proposed hypotheses, we see that data availability, both financial and non-financial, is positively associated with credit risk. To corroborate our results, we have gone a step further by first calculating Altman-Z score for our sample and seeing whether credit risk calculated with the Bharath and Shumway formula varies depending on if we have enough data to obtain Altman-Z score.

Altman-Z score is another measure of risk, that was "*introduced as a way of predicting the probability that a company would collapse in the next two years*". (CFI Team, 2022) The formula to calculate it is the following:

$$Z - score = 1{,}2\left(\frac{Working\ capital}{Total\ assets}\right) + 1{,}4\left(\frac{Retained\ earnings}{Total\ assets}\right) + 3{,}3\left(\frac{Earnings\ before\ interest\ and\ taxes}{Total\ assets}\right) + 0{,}6\left(\frac{Market\ value\ of\ equity}{Total\ liabilities}\right) + 1.0\left(\frac{Sales}{Total\ assets}\right)$$

Once we obtained Altman-Z scores for the sample, we created a Dummy variable so that the blank cells, those for which Altman Z-score could not be calculated because there was not enough data, taking value 0 and 1 otherwise. Taking this into account, we then tested for homogeneity of variances and equality of means and obtained the following table:

Table 16. Credit risk variance and mean test for high-low Altman-Z score

	Statistic	Altman-Z availability	Altman-Z no availability	Difference
Credit risk (BS)	Variance	0,0273	0,03255	-0,0052***
	Mean	0,0429	0,05253	-0,0097***

The table shows the tests for equal variances and means for the Credit Risk variable we there is data availability to obtain Altman-Z score or not. Altman-Z score availability is the independent variable, and Credit Risk is the dependent variable ***, **, and * indicates statistical significance at 1 percent, 5 percent, and 10 percent levels. High availability (dummy) takes value 1 if company discloses information for more than 10 variables and 0 otherwise.

Firstly, equal population variances could not be assumed for those firms that presented enough data for us to calculate their Altman-Z score and those that did not, with a level of significance of 1%. As for equality of means, we find a negative association between credit risk calculated with Bharath and Shumway and having data to measure Altman-Z score, with a level of significance of 1%.

In conclusion, investors should interpret the non-ability to obtain Altman-Z score for a business as a negative sign related to their credit risk.

5. CONCLUSIONS

As transparency has become increasingly important nowadays, and conforming data availability a big part of it, the objective of this paper was to examine whether data availability was in any way related with a firm's performance and its default risk. Moreover, because of the importance ESG factors have gained in the last few years as stakeholders pressure firms to report on them, our aim was also to find if there was any relationship between firms' performance and default risk and their ESG disclosure.

We proposed our hypotheses from a "*managerial bad news hoarding*" perspective (Da Silva, 2022) together with a "*signaling theory*" perspective (Jullobol and Sartmool, 2015). This means that firms' managers would try to hide the bad news as much as possible, while they would be eager to disclose the good ones as fast as possible to differentiate themselves from others. In the end, companies are made up by people, and their decisions are conditioned by human behavior. And in the same way that we post the trip of our dreams on social media and hide our bad days or state our best skills in our resume but not our flaws, it would make sense for firms to quickly disclose their successes to appeal to investors and hide their failures as long as possible.

This paper has allowed us to confirm a significant positive relationship between Return on Assets and data availability and a significant negative relationship between Return on Equity and data availability and credit risk and data availability. Moreover, we found a significant positive relationship between financial disclosure and ESG disclosure and between financial performance and ESG disclosure, and a significant negative relationship between default risk and ESG disclosure.

Taking these results into account, it would be smart for investors to interpret a higher amount of data availability as a positive sign in terms of a firm's ability to pay back a debt. So, in a situation where an investor had to choose between two apparently identical US firms, choosing the one that was disclosing more information, both financial and non-financial, would be the way to go.

On another note, a limitation of our analysis and an opportunity for future research is to analyze each of the ESG factors separately: "Environmental", "Social" and "Governance". Moreover, we have also said that data availability is just a part of transparency. This is because transparency worries not only about quantity but quality. And quality refers to how the story is told, how easy or complex the report is to read: a report could be longer and provide with more information but do it in a misleading way. We realize that our analysis does not provide insights about the quality side of transparency, which is quite important. We consider this issue a limitation of the current analysis and an opportunity for future research. In effect, constructing a transparency measure and examining its relationship with default risk could prove very useful for investors. Another

As a way to put an end to this paper, we want to highlight that increasing disclosure is the first step to achieve transparency and to build trust among stakeholders, which is not an easy task nowadays. And trust matters. As stated in the PwC Annual Report 2022: "*Trust is the hallmark of high-performing companies, the glue that binds cohesive societies, and a driver of shared prosperity. And trust can't be bought. It must be earned through every interaction, every experience, every relationship and every outcome delivered.*"

REFERENCES

Abinzano, I., Gonzalez-Urteaga, A., Muga, L., & Sanchez, S. (2020) Performance of default-risk measures: the sample matters. *Journal of Banking and Finance*, 120, 105959. https://doi.org/10.1016/j.jbankfin.2020.105959

Adeneye, Y. and Ahmed. M (2015), Corporate social responsibility and company performance, *Journal of Business Studies Quarterly* 2152-1034. 7. 151-166.

Adiloğlu, B. and Vuran, B. (2017). The importance of EBIT- EBITDA disclosure in annual reports: A comparison from Turkey. *International Journal of Social Sciences and Education Research.* 3. 397-397. 10.24289/ijsser.284252.

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

Anderson, A. and Folkare, E. (2015). Company characteristics and voluntary disclosure of intellectual capital – A study on Swedish listed companies. *Jönköping University*

Balakrishnan, K., B. Billings, M., T. Kelly, B. and Ljungqvist A. (2014), Shaping liquidity: On the causal effects of voluntary disclosure. *Journal of Finance* 69, 2237–2278.

Bharath, S.T., & Shumway, T. (2008) Forecasting Default with the Merton Distance to Default Model. *The Review of Financial Studies, 21*(3), 1339-1369.

Botosan, C., A. and A. Plumlee, M. (2002), A Re-Examination of Disclosure Level and the Expected Cost of Equity Capital, *Journal of Accounting Research* 40, 21–40.

Buallay, A. (2018) Is sustainability reporting (ESG) associated with performance? Evidence from the European banking sector, *Management of Environmental Quality: An International Journal*, https://doi.org/10.1108/MEQ-12-2017-0149

Campbell, D., Shrives, P., and Bohmbach-Saager, H. (2001). Voluntary disclosure of Mission Statements in Corporate Annual Reports: Signaling What and to Whom? *Business and Society Review* 106:1, 65-87.

CFI Team (2022). Altman's Z-Score Model. *Corporate Finance Institute.* Retrieved from https://corporatefinanceinstitute.com/resources/commercial-lending/altmans-z-score-model/

Da Silva, P. (2022). Crash risk and ESG disclosure. *Borsa Istanbul Review*. Volume 22, Issue 4, Pages 794-811. https://doi.org/10.1016/j.bir.2022.04.001

Gharghori, P, Chan, H., & Faff, R. (2006) Investigating the performance of alternative default-risk models: Option-based versus accounting-based approaches. *Australian Journal of Management*, 31, 2, 207-234.

Gow, I., Taylor, D. and Verrechia, R. (2011). Disclosure and the Cost of Capital: Evidence of Information Complementarities

Hillegeist, S.A., Keating, E.K., Cram, D.P., & Lundsted, K.G. (2004) Assessing the Probability of Bankruptcy. *Review of Accounting Studies*, 9, 5-34.

Hilscher, J., & Wilson, M. (2017). Credit ratings and credit risk: Is one measure enough? *Management Science*, 63(10), 3414-3437.

Kealhofer, S. (2003) Quantifying Credit Risk I: Default Prediction. *Financial Analysts* Journal, 59, 1, 30- 44.

Holanders, B. (2013) Financial reporting transparency and Peer selection for RPE. *Tilburg School of Economics and Management*

Jullobol N. and Soibuppha S. (2015). The effect of firm performance on voluntary disclosure in annual reports: a case study of technology industry in the stock exchange of Thailand. *RMUTT Global Business and Economics Review*

Kim, S.,Kraft, P. and Ryan, S. (2013). Financial Statement Comparability and Credit Risk. *Rev Account Stud* 18, 783–823.

Lambert, R., Leuz, C., & Verrecchia, R. E. (2007). Accounting information, disclosure, and the cost of capital. *Journal of Accounting Research*, 45(2), 385–420.

Leuz, C. and R. Verrecchia (2000), The Economic Consequences of Increased Disclosure, *Journal of Accounting Research* 38, 91–124.

Li, H., Zhang X. and Zhao, Y. (2022) ESG and Firm's Default Risk. *Finance Research Letters*, Volume 47, Part B. https://doi.org/10.1016/j.frl.2022.102713.

McClure, B. (2021) Financial Reporting: The Importance of Corporate Transparency. *Investopedia*

Melloni, G., Caglio, A., and Perego, P. (2017). Saying more with less? Disclosure conciseness, completeness and balance in Integrated Reports. *Journal of Accounting and*

Public Policy Volume 36, Issue 3, Pages 220-238. http://dx.doi.org/10.1016/j.jaccpubpol.2017.03.001

PwC Annual Report (2022). Available at pwc.com/annualreview

Schön, W. (2006), Corporate Disclosure in a Competitive Environment - the Ecj's Axel Springer Case and the Quest for a European Framework for Mandatory Disclosure, *ECGI - Law Working Paper* No. 55/2006.

Seran, A.M.I. (2022). Analysis Of The Effect Of Firm Characteristics On The Completeness Of Public Company's Voluntary Disclosure. *Journal of Business and Behavioural Entrepreneurship*, 6(1), 51-68.

Tamimi, N. and Sebastianelli, R. (2017). Transparency among S & P 500 companies: an analysis of ESG disclosure scores. *Management Decision*. Vol. 55 No. 8, pp. 1660-1680.

Truong, L., Thai L., and Swint H. (2022). The Influence of InformationTransparency and Disclosure on the Value of Listed Companies: Evidence from Vietnam. Journal of Risk and Financial Management 15: 345.

Yu, F. (2003). Accounting Transparency and the Term Structure of Credit Spreads. Journal of Financial Economics. Vol.75 No.1, pp.30-47