

# RESEARCH PROPOSAL FOR MASTER THESIS

CHIA, Jeremy Jerome

M.Sc. Big Data and Business Analytics (2022 Promotion)

---

## Proposed Topic:

A Natural Language Processing (Topical and Sentiment) Analysis of Sustainability Reports and its Relationship with Corporate Financial Performance: Empirical Evidence from Singapore's Listed Companies

---

## Introduction

Non-financial information is increasingly important to stakeholders in assessing a company's performance. This morphs the classical view that the shareholders exclusively seek to maximise returns. For example, the *triple-bottom-line* – proposed in *Elkington (1999)* proposed that companies were to focus on people and the planet and profits. *Tsang (1998)* observed that several listed companies in Singapore's banking, food and beverages, and hotel industries were making voluntary disclosures on their social corporate responsibility efforts. Such disclosures were often a section in the annual report. These are centred on human resources and community involvement efforts, at 5.3% of all sentences in the annual report. However, environment-related disclosures were scarce – they made up only 0.24% of total sentences. These results were representative of the broader social focus in Singapore in the 1990s on fostering social cohesion. Since that initial study, other similar studies have revealed statistically significant positive relationships between Corporate Social Responsibility (CSR) disclosures and financial performance (*Khavesh et al., 2012*).

Sustainability reports have become commonplace for companies to communicate their non-financial performance and initiatives to their stakeholders. These reports are now part of mandatory disclosures for listed companies ("issuers") in Singapore. From financial periods ending on or after 31 December 2017, companies listed on the Singapore Exchange (SGX) were required to issue sustainability reports. Practice Note 7.6 stipulated that companies shall describe their sustainability practices on a "comply or explain" basis (*Chia and Wee, 2021*). The Exchange proposed that firms may use widely adopted frameworks, such as the Global Reporting Initiative (GRI) and the Sustainability Accounting Standards Board (SASB), as a basis for sustainability reporting (*Loh et al., 2017*). The use of standard frameworks allows for comparability. This also avoids the need for companies to use costly sustainability initiatives to reduce information asymmetry. (Connelly, Ketchen & Slater, 2011)

Following the "Sustainability Reporting Review 2021" by the Task Force on Climate-related Financial Disclosures (TCFD), SGX has moved to strengthen its sustainability reporting requirements. From 1 January 2022, issuers must perform 'climate reporting'. This will become mandatory for issuers in the (a) financial industry, (b) agriculture, food and forest products industry, (c) energy industry, (d) materials and buildings industry, and (e) transportation industry. A "mandatory" requirement departs from the prior position of "comply or explain". This shift in regulation signals the Exchange's initiative to combat climate change and its value on non-financial information such as sustainability reports, allowing stakeholders to assess better the issuer's financial prospects and quality of management (*Singapore Exchange, 2022*).

There has been a heightened emphasis on climate change and increasing prominence of social issues in recent years. Regulatory and social developments may result in varied impacts on sustainability disclosures on corporate financial performance (CFP) compared to findings in prior studies. For example, *Loh et al., 2017* studied the relationship between CSR disclosures and CFP when SGX instituted a “comply or explain” regime. This was a change from its previous position – encouraged but voluntary disclosures. At this juncture where SGX is strengthening its requirements – making CSR disclosures mandatory for TCFD-identified industries, this study will examine if the relationship for Singapore-listed companies is materially different.

There were studies conducted using empirical evidence from listed companies in Singapore (*Tsang, 1998; Loh et al., 2017; Wahyuningrum et al., 2021*), as well as other jurisdictions – Malaysia (*Kasbun et al., 2017*), India (*Aggarwal, 2013-a; Motwani and Pandya, 2016*), Brazil (*Ching et al., 2017*), Australia (*Jones et al., 2007*). These studies regressed financial metrics against sustainability indicators at a point in time. In determining inputs for sustainability indicators, these studies used simple, self-developed scoring methods to assess the quantity and quality of disclosures. These approaches will be discussed in the subsequent section.

The proposed study will adapt prior approaches and proposed new methods in assessing the relationship between CSR disclosures and CFP. Two Natural Language Processing (NLP) techniques are adopted to analyse annual reports – topical modelling (*Rehurek and Sojka, 2011*) and sentiment analysis (*Pang et al., 2002*). These approaches provide a robust yet different dimension of text mining sustainability reports. Input variables from this analysis are regressed against indicators of CFP, such as Return on Equity (ROE) and Return on Assets (ROA), amongst others. The approach will be discussed further in the Methodology section.

The research questions are, in the context of Singapore-listed companies:

1. What are the most prominent topics in sustainability reports from an exploratory topic modelling analysis? Will this resemble the disclosure themes in frameworks like the GRI and SASB?
2. How do these topics identified in (1) relate to CFP?
3. How does the sentiment score of a sustainability report relate to corporate financial performance?
4. Are there material differences in the relationship between CSR disclosures and CFP at this point, vis-à-vis prior studies conducted by *Tsang, 1998; Loh et al., 2017* and *Wahyuningrum et al., 2021*?

## 2. Literature Review

### *Theoretical Discussions*

*Sustainability reporting* is a complex phenomenon that a single theory cannot explain (Cormier *et al.*, 2005; Tagesson *et al.*, 2009). Therefore, unifying all empirical findings within one theoretical framework remains a challenge. Nevertheless, a combination of several theories – the legitimacy, signalling, and agency theory – aids an understanding of the theoretical relationship between CSR disclosures and financial performance.

The legitimacy theory suggests that it is necessary to achieve society's approval for the company to survive. Companies need to act in congruence with society to uphold their business activities. (O'Donovan, 2002). Companies are encouraged to seek legitimacy through disclosures (Cho and Patten, 2007). CSR disclosures provide information that legitimises the company's behaviour to influence stakeholders' and, eventually, society's perceptions about the company, resulting in higher firm value and better financial performance (Hooghiemstra, 2002). These are exemplified through the topics companies choose to report.

Signalling theory underpins the notion that an agent (company executives) conveys information to the principal (stakeholders). In the context of sustainability, investors and consumers have difficulties in determining which firms are genuinely committed to sustainability. (Connelly *et al.*, 2011) Companies that disclose environmental issues signal that they are engaged in proactive environmental strategy. There is an incentive to inform principals by voluntarily disclosing more, as positive signals are more appealing (Clarkson *et al.*, 2008). Sentiment analysis of sustainability reports may reveal insights into the relationship between the sentiments expressed in these reports and financial performance.

In tandem with agency theory, voluntary disclosure of firms, mainly on social and environmental aspects, is a means to reduce current or potential agency costs that may occur in the form of legislation and regulation (Galani *et al.*, 2012). This would close the gap between the misalignment of objectives between principals and their agents. By way of reducing information asymmetry, sustainability reporting would have positive effects on a firm's perception.

### *Conceptual Frameworks*

Conceptual frameworks expanded on the classical theoretical foundations in providing a contextual understanding of this phenomenon. Richardson *et al.* (1999) proposes a model to explain the capital market impacts of CSR, shown in *Figure 1*. In addition, Reddy and Gordon (2010) present a conceptual framework to explain the causality of corporate environmental efforts with firm value, shown in *Figure 2*.

These conceptual frameworks suggest that CSR efforts will impact firm value. For example, a key intermediary in the process is signalling through disclosures. Communicating their efforts through disclosures would affect perceptions of risk, performance, and management, leading to effects on CFP.

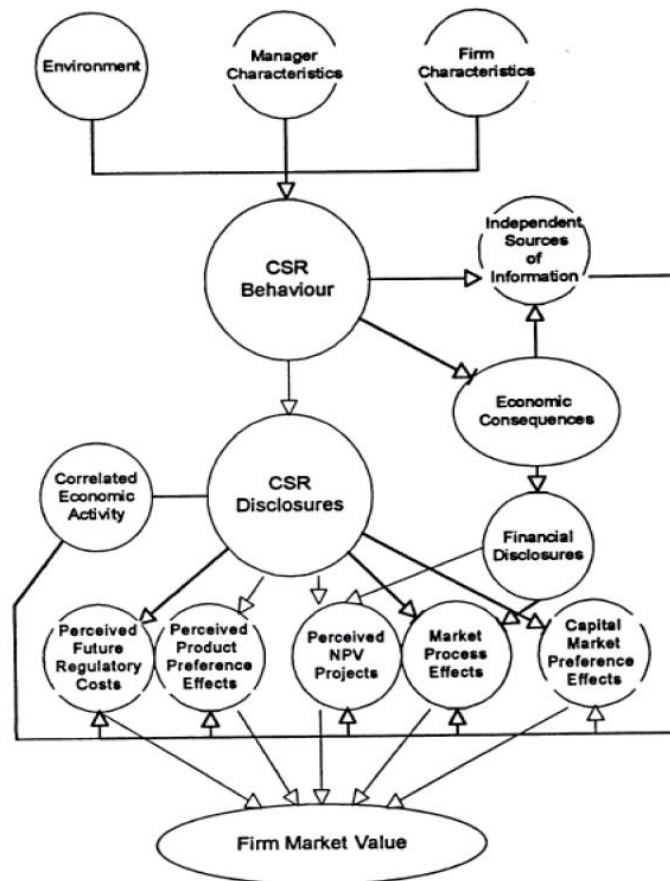


Figure 1: A model of the capital market impacts of corporate social responsibility (Richardson et al., 1999)

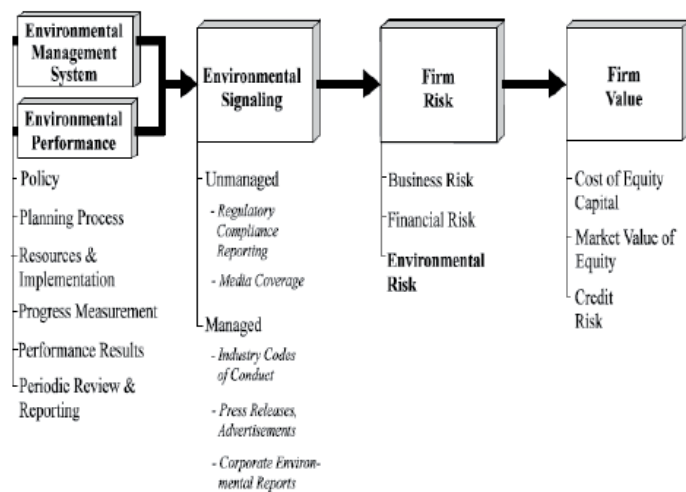


Figure 2: Conceptual Model Linking Corporate Environmental Management and Performance with Firm Value (Reddy and Gordon ,2010)

### *Prior Approaches*

Prior studies conducted on this topic used empirical evidence to test the statistical significance of the relationship between CSR disclosures and CFP. These have minute variances. Some studies chose to regress listed firms at a point in time to determine if more disclosures in the year had a positive relationship with financial performance (Kasbun *et al.*, 2017; Aggarwal, 2013-a; Motwani and Pandya, 2016; Ching *et al.*, 2017; Jones *et al.*, 2007). Others conducted a longitudinal study comparing companies' performance across years (Tsang, 1998; Wahyuningrum *et al.*, 2021). In longitudinal studies, the authors sought to examine if an improvement in CSR disclosures resulted in a similar improvement in CFP.

The factors used to determine sustainability disclosures and performance were varied. Loh *et al.* (2017) used a measurement scheme developed by the ASEAN CSR Network, which proposes 23 criteria in the scheme over governance, economic, environmental, and social indicators. Wahyuningrum *et al.* (2021) used an abridged form of content analysis to assign scores based on the number of disclosures – the number of sentences number of pages. Kasbun *et al.* (2016) similarly used the number of sentences to indicate the quantity of activity reported. Aggarwal (2013-a) used scores extracted from private databases provided by Bloomberg and Thomson Reuters Eikon. The majority of these metrics can often be superficial and do not actively examine the text of such CSR disclosures, which is proposed in this study.

Financial metrics used were based on generally-available information. For example, figures were retrieved from audited financial reports of the same year. Standard metrics used to assess performance were the Return on Assets (ROA), Return on Equity (ROE), Profit before Tax (PBT), employed in several studies (Aggarwal, 2013-a; Ching *et al.*, 2017; Kasbun *et al.*, 2016). Jones *et al.*, 2007 sought to examine the effect on abnormal stock returns.

Studies examined used regression analysis in testing the statistical significance. Loh *et al.* (2017) used the Ohlson model (Ohlson, 1995) as a baseline model and added proposed sustainability indices to determine the improved significance of the models to determine the explanatory effect of sustainability variables included. Several others opted to use each financial indicator as a dependent variable iteratively and sustainability factors as an independent variable. Firm size, a function of the total assets, was also considered in the model.

Some studies sought to distinguish results by industry and other characteristics. For example, Ching *et al.* (2017) looked only at companies that were determined to have the best level of disclosure – for example, companies classified as Application Level “A” under the GRI. Loh *et al.* (2017) differentiated the results by industry and the status. They considered whether being a Government-Linked Company or a family-owned business should have significant differences. Distinguishing companies with these characteristics is another potential approach. Examining the distinction between different industries is worth considering, given that the TCFD's recommendations have only singled out several industries for disclosures.

Several elements used by prior studies in a Singapore context will be retained to provide a basis for comparison. The choice of financial metrics will be one of them. This study, however, will seek to propose a robust method of examining sustainability disclosures by examining the text of the sustainability reports.

## **Methodology**

The proposed study is a longitudinal analysis of changes in companies' CSR disclosures for financial periods ending in 2016 to 2021 and their relation with financial performance.

### *Data Used*

Sustainability reports and annual reports will be retrieved from the SGX website. Topical modelling (*Rehurek and Sojka, 2011*) and sentiment analysis (*Pang et al., 2002*) will be used to analyse the text in the sustainability reports. Financial performances indicators will be extracted from information available from audited financial statements within the annual reports.

### *Models*

In topical modelling, text within the sustainability reports will be analysed by clustering to determine the most critical topics. A widely-adopted framework, the GRI, sustainability reporting typically encompasses three material topics – economic topics, environmental, and social. Governance is also often associated with sustainability reporting. A topical analysis provides insights on the actual topics disclosed and if there are any trends in a local specific context in Singapore. With the topics identified, further processing can be performed to derive independent variables, such as the proportion of disclosures, the extent of disclosures, and the use of specific words.

The same corpus of text within the sustainability reports will be analysed for sentiment analysis to derive overall “sentiment” scores. A starting point for such analysis could be using the Valence Aware Dictionary, and Sentiment Reasoner (VADER) proposed in *Bird et al. (2009)*. Other dictionaries can also be explored. These scores would result in an input variable.

A multivariate regression will be performed to test the significance of the model, similar to other studies.

## **Potential Implications**

This study is exploratory. It seeks to determine if using Natural Language Processing will result in a method for better explaining the relationship between CSR disclosures and CFP. If the results are significant, the dual-pronged approach of topic modelling and sentiment analysis can be adopted for prospective studies.

It is expected that there will be topics that present a more positive relationship. This could be of interest to sustainability communications professionals and the company's management. Categories of CSR impact which evidence a more positive relationship may be of interest to these stakeholders. Companies can employ this information to increase the number of initiatives and perform more disclosures to boost their financial performance. These boost financial performance and other objectives of the triple bottom line concurrently.

It would also be helpful to take stock of the impact of the change in SGX's regulations. For example, the Exchange's move to make climate reporting mandatory for TCFD-identified industries is yet another milestone in developing Singapore's sustainability reporting requirements. As Singapore strives to continue to present itself as a responsible member of the global community on green-related issues, sustainability disclosures are an essential component of this strategy. Understanding and assessing the effects would encourage but also inform further initiatives.

## Bibliography

Aggarwal, P. (2013a). Impact of sustainability performance of company on its financial performance: A study of listed Indian companies. *Global Journal of Management and Business Research (C: Finance)* Volume, 13.

Aggarwal, P. (2013b). Sustainability reporting and its impact on corporate financial performance: A literature review. *Indian Journal of Commerce and Management Studies*, 4(3), 51

Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python: analyzing text with the natural language toolkit. " O'Reilly Media, Inc."

Chia K. H., Wee, E. (2021, December 28). Enhanced disclosures on climate-related information & board diversity policy for SGX-listed companies. Lexology. Retrieved from <https://www.lexology.com/library/detail.aspx?g=1b4ea3d2-1330-4f35-8e5f-447f181e3e64> on 2022, April 13

Ching, H. Y., Gerab, F., & Toste, T. H. (2017). The quality of sustainability reports and corporate financial performance: Evidence from Brazilian listed companies. *Sage Open*, 7(2), 2158244017712027.

Cho, C. H., & Patten, D. M. (2007). The role of environmental disclosures as tools of legitimacy: A research note. *Accounting, organizations and society*, 32(7-8), 639-647.

Clarkson, P. M., Li, Y., Richardson, G. D., & Vasvari, F. P. (2008). Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. *Accounting, organizations and society*, 33(4-5), 303-327.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of management*, 37(1), 39-67.

Cormier, D., Magnan, M., & Van Velthoven, B. (2005). Environmental disclosure quality in large German companies: economic incentives, public pressures or institutional conditions?. *European accounting review*, 14(1), 3-39.

Elkington, J., & Rowlands, I. H. (1999). Cannibals with forks: The triple bottom line of 21st century business. *Alternatives Journal*, 25(4), 42.

Galani, D., Gravas, E., & Stavropoulos, A. (2012). Company characteristics and environmental policy. *Business Strategy and the Environment*, 21(4), 236-247.

Hooghiemstra, R. (2000). Corporate communication and impression management—new perspectives why companies engage in corporate social reporting. *Journal of business ethics*, 27(1), 55-68.

O'Donovan, G. (2002). Environmental disclosures in the annual report: Extending the applicability and predictive power of legitimacy theory. *Accounting, Auditing & Accountability Journal*.

Jones, S., Frost, G., Loftus, J., & Van Der Laan, S. (2007). An empirical examination of the market returns and financial performance of entities engaged in sustainability reporting. *Australian accounting review*, 17(41), 78-87.

Khavesh, A., Nikhasemi, S. R., Haque, A., & Yousefi, A. (2012). Voluntary sustainability disclosure, revenue, and shareholders wealth-a perspective from Singaporean companies. *Business Management Dynamics*, 1(9), 06-12.

Kasbun, N. F., Teh, B. H., & San Ong, T. (2016). Sustainability reporting and financial performance of Malaysian public listed companies. *Institutions and Economies*, 78-93.

Loh, L., Thomas, T., & Wang, Y. (2017). Sustainability reporting and firm value: Evidence from Singapore-listed companies. *Sustainability*, 9(11), 2112.

Motwani, S. S., & Pandya, H. B. (2016). Evaluating the impact of sustainability reporting on financial performance of selected Indian companies. *International Journal of Research in IT and Management*, 6(2), 14-23.

Oncioiu, I., Petrescu, A. G., Bîlcan, F. R., Petrescu, M., Popescu, D. M., & Anghel, E. (2020). Corporate Sustainability Reporting and Financial Performance. *Sustainability*, 12(10), 4297.

Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. *arXiv preprint cs/0205070*.

Reddy, K., & Gordon, L. (2010). The effect of sustainability reporting on financial performance: An empirical study using listed companies.

Rehurek, R., & Sojka, P. (2011). Gensim–python framework for vector space modelling. NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic, 3(2).

Richardson, A. J., Welker, M., & Hutchinson, I. R. (1999). Managing capital market reactions to corporate social responsibility. *International Journal of Management Reviews*, 1(1), 17-43.

Singapore Exchange. (2022). *Sustainability reporting*. Singapore Exchange - Singapore Exchange (SGX). Retrieved from <https://www.sgx.com/regulation/sustainability-reporting> on 2022, April 13

Tagesson, T., Blank, V., Broberg, P., & Collin, S. O. (2009). What explains the extent and content of social and environmental disclosures on corporate websites: a study of social and environmental reporting in Swedish listed corporations. *Corporate social responsibility and environmental management*, 16(6), 352-364.

Tsang, E. W. (1998). A longitudinal study of corporate social reporting in Singapore: The case of the banking, food and beverages and hotel industries. *Accounting, Auditing & Accountability Journal*.

Wahyuningrum, I. F. S., Oktavilia, S., Putri, N., Solikhah, B., Djajadikerta, H., & Tjahjaningsih, E. (2021). Company financial performance, company characteristics, and environmental disclosure: evidence from Singapore. In *IOP Conference Series: Earth and Environmental Science* (Vol. 623, No. 1, p. 012065). IOP Publishing.