26 August 2025 ·

Qeios

Peer Review

Review of: "Greening the Bottom Line: Exploring the Impact of Sustainability Disclosure on Financial Performance"

Marius Muller[1]

1. Construction Economics and Management, University of Witwatersrand, South Africa

Introduction

The study answers the question of what the relationship is between the quality of sustainability disclosure and the mining industry's financial success in South Africa. There are major methodological problems with the manuscript that make the findings less reliable. The problems range from measurement tools that are not in line with current standards to the economic model not having any diagnostic tests.

Data and Measurements

There are some problems with the study's core statistics that make it hard to trust it:

The measurement tool for the study does not work with the reporting standards that were in place for most of the study period. It is based on the GRI G3 standards, which were finalized in 2011. But between 2012 and 2021, the official GRI framework changed a lot. In 2013, the materiality-focused G4 standards were released, and in 2016, the modern GRI Standards were released. The study compares all companies to a standard that is becoming less useful by using a static G3-based index. This means it is likely that the tool punished businesses for using the latest best practices. This adds a measurement error that makes the whole sample look questionable.

A 0/1 ("not met"/"met") measure was used to come up with the study's sustainability quality score. This method is too simple because it cannot tell the difference between basic and detailed information. It may measure compliance rather than real "quality" because it treats a company that gives one vague sentence about a topic the same as one that gives a detailed, publicly audited analysis.

 doi.org/10.32388/K1T66V

There is no information in the paper about whether the scores of the 324 company reports were checked by a second coder. However, since there are two writers, this could have happened, but it is not made clear. It is hard to be sure that the scores were given consistently and without subjective bias without a reported measure of inter-coder reliability. This is a major threat to the validity of the measurement.

For missing values, mean imputation was used, and winsorizing was used to deal with outliers. However, the paper does not include robustness checks to show that the main findings are still true when these values are added to the analysis. For instance, the range for some variables is very large. For example, the liquidity spread is 209.66 (0.120 min - 209.780 max), and its mean is 6.812. An explanation or robustness check to support the same would be of value.

Econometric Model

There are a few problems with the model that was used to test the hypotheses. When the paper talks about how "appropriate" a fixed effects model was, it does not show any Hausman test results to back up this choice. There is also no reference to a test for heteroskedasticity. Table 6's model has a term indicated by AR(1) lag to fix serial correlation, but it does not show a test to find this problem in the first place. Lastly, the model does not check for cross-sectional dependence, so it does not consider industry-wide shocks like changes in product prices that could affect all firms at the same time and make the results hard to understand.

Additionally, the R-squared value given by the model is 0.961. This value is unusual and could possibly suggest that the model may be "overfitted or suffering from multi-collinearity." However, this may be completely plausible, but no evidence is provided to support this. The paper also finds that the quality of the reporting affects how well a business does, but it does not take into account the strong chance of a "chicken-and-egg" problem. It is possible that businesses that make more money can afford to produce better reports. Endogeneity is not addressed either.

General

The way the introduction and literature parts are written makes them sound more like separate statements than a well-thought-out argument or discussion. Use of a different citation style, like Harvard or APA, would be more common and make the text easier to read. There are a lot of conflicting studies on the subject of this paper, as shown in Table 1's literature review. The paper does not try to synthesize these results or explain why there might be such conflicts. It would be good to review and address the pointers highlighted above before submitting for further review.

Declarations

Potential competing interests: No potential competing interests to declare.

doi.org/10.32388/K1T66V