Big data analytics and environmental performance: The moderating role of internationalization

1. Introduction

In the contemporary business landscape, organizations are increasingly leveraging big data analytics (BDA) to gain insights into their operations, enhance decision-making processes, and achieve competitive advantages (Sivarajah *et al.*, 2017). The advance in data utilization techniques, underpinned by artificial intelligence and machine learning (Hernández *et al.*, 2018), has enabled the collection and analysis of data characterized by immense volume, diverse variety, real-time availability, quality, and the ultimate goal of deriving value and insights for organizations, encapsulated by the concept of the five Vs (volume, variety, velocity, veracity, and value) (Emani, Cullot and Nicolle, 2015; Mikalef *et al.*, 2019). Employing these techniques is anticipated to strengthen the credibility and effectiveness of analyses and insights into descriptive, prescriptive, and predictive knowledge, offering enhanced support for decision-making within organizations (Phillips-Wren and Hoskisson, 2015; Wang et al., 2018).

While much effort has been devoted to investigating the impact of big data on financial performance (Akter et al., 2016; Boso et al., 2013) and business entry (Wei, Li and Wang, 2024), there has been an emerging literature concerning the impact of BDA on environmental performance, the majority of which either propose theoretical frameworks (Queiroz, 2018; Roman Pais Seles et al., 2018; Song et al., 2017; Song et al., 2018; Wu et al., 2017) or rely on subjective measures, such as surveys or case analyses (Belhadi et al., 2020; Calic and Ghasemaghaei, 2021; Cheng and Liu, 2018; Dubey et al., 2019). Existing studies have yet to empirically establish the causal link between BDA and environmental performance, primarily due to a reliance on perceived benefits, which may not accurately reflect the actual impact of BDA initiatives within organizations and could lead to a reverse relationship. Additionally, there is a crucial need to investigate whether ongoing BDA initiatives within organizations either enhance or diminish tangible business benefits over time (Zhu, Dong and Luo, 2021). Belhadi et al. (2020) call for further longitudinal research to empirically validate existing theories linking BDA to environmental performance, an aspect currently unexplored in literature, thus leaving a gap in understanding the environmental value of BDA efforts.

Simultaneously, international business studies maintain that "going-global" strategies hold significant importance in fostering responsible business practices (Bansal, 2005; Aguilera-Caracuel, Hurtado-Torres and Aragón-Correa, 2012). Consequently, escalating environmental concerns pose substantial challenges for internationalizing firms due to divergence among countries regarding their willingness and capacity to address climate and environmental issues (Bansal, 2005; Lubinski and Wadhwani, 2020). Barbosa et al. (2022) indicate that the domestic market alone might not ensure sustainable performance; instead, other types of influence, such as engaging in internationalization activities, may be necessary to achieve higher sustainable and environmental outcomes. Further, previous studies have also developed preliminary framework linking BDA and internationalization (Cheng et al., 2020; Dam et al., 2019; Gnizy, 2019). Hence, it can be inferred that internationalization, might hold theoretical and practical significance in influencing the BDA - environmental performance relationship for internationalizing firms. However, there is a dearth of articles investigating the integration of these three concepts (Ahi *et al.*, 2022; Ocelík, Kolk and Ciulli, 2023), highlighting a need for a multidisciplinary framework to clarify existing knowledge gaps regarding the convergence of BDA, internationalization, and environmental sustainability.

To fill such research voids, we concentrate on internationalizing firms - those expanding beyond their domestic borders, typically through exporting, foreign direct investment, or establishing international partnerships or subsidiaries (Blomstermo *et al.*, 2004) - which are confronting new challenges that compel them to reevaluate their approach to leverage digitalization and sustainability (George and Schillebeeckx, 2022) to investigate the impact of BDA implementation on environmental performance and explore how aspects of internationalization, namely internationalization degree, speed, scope and rhythm, moderate this relationship. Drawing from the organizational learning theory, this paper, based on the panel data from US Fortune 500 listed companies from 2010 to 2022, confirms the positive impact of BDA utilization on environmental performance and further explains the moderating effect of internationalization.

Our study enhances the literature on BDA and environmental performance in multiple dimensions. First, we contribute to the understanding of the BDA-environmental performance link, addressing the lack of longitudinal analysis by examining data from 2010 to 2021. Our results not only corroborate earlier findings (Belhadi et al., 2020; Calic and Ghasemaghaei, 2021; Cheng and Liu, 2018; Dubey et al., 2019) but also demonstrate the causal relationship between BDA and environmental performance. Secondly, we apply organizational learning theory to BDA, an underexplored area, expanding its application beyond traditional organizational contexts (i.e., Ali, 2021; Asimakopoulos et al., 2020). Third, unlike prior studies that focus on individual relationships between BDA, environmental performance, and internationalization, our study stands out as one of the pioneering endeavors to empirically explore the integrated relationship among these constructs, demonstrating their combined impact on environmental performance.

This paper is structured as follows: Section 2 outlines the research hypotheses, while Section 3 discusses data sources, variables, and methodology. Section 4 presents empirical results, followed by a robustness test to validate our findings. Finally, the conclusion highlights the study's significance and suggests avenues for future research.

2. Theoretical Analysis and Research Hypothesis

Organizational learning theory suggests that a company’s performance is largely contingent upon its systematic pursuit of information, which enhances its ability to adapt and innovate in various areas, including processes, products, and services (Huber, 1991). From the standpoint of organizational learning theory, the dimensions of big data—such as volume, velocity, and variety—enhance the feasibility of organizational learning in companies. In the context of environmental sustainability, organizational learning theory offers a theoretical lens to understand how utilization and processing of rich information from BDA can lead to increased environmental performance.

Recent literature has highlighted that big data holds significant potential to impact environmental studies (Sarker et al., 2020; Song, et al., 2017). BDA serves as an organization’s strategy to facilitate the acquisition, organization, and assessment of extensive data from diverse sources to identify patterns and trends (Dubey *et al.*, 2020; Riggs *et al.*, 2023). These insights can then be utilized to strategically prioritize resource efficiency and sustainability initiatives, such as waste reduction, material reuse, and product recycling (Gupta *et al.*, 2019; Kristoffersen *et al.*, 2021). A business report by 3p Contributors (2016) illustrates how Pirelli, the world’s fifth-largest tire manufacturer, collaborated with SAP to use real-time big data for efficient inventory management, enabling Pirelli to prevent tire disposal in landfills and reduce greenhouse gas emissions by proactively addressing inventory challenges. Beside these anecdotal examples, prior research also offers similar evidence regarding the types of environmental benefits linked to environmental performance through BDA. In a separate study, Koseleva and Ropaite (2017) underscore the utility of big data in enhancing energy-efficient practices. Meanwhile, Song et al. (2017) delve deeper into the role of big data in supporting organizational objectives to enhance social and environmental sustainability. These discussions provide theoretical and initial limited empirical evidence, although comprehensive longitudinal validation for such benefits derived from BDA is lacking. Hence, we hypothesize that:

Hypothesis 1. *BDA utilization has a positive impact on environmental performance.*

González-Benito and González-Benito (2006) propose a theoretical framework suggesting that the degree of globalization within firms influences environmental proactivity through various channels, e.g., cross-sector knowledge exchange, the introduction of new environmental policies to meet the sustainability requirements of importing economies, and the creation of opportunities to access additional resources. On the other hand, as firms expand internationally, they encounter heightened levels of operational complexity and uncertainty (Kogut, 1989). This entails navigating diverse external environments with varying cultures, regulations, and stakeholder demands diverging from their domestic context. Given these complexities, we argue that internationalization positively moderates the impact of BDA utilization on environmental performance through two primary channels.

From the demand side, stakeholder theory (Freeman and Liedtka, 1997) emphasizes a firm's success depending on responsiveness to stakeholders; yet, differing perspectives across national contexts (Asmussen and Fosfuri, 2019) and institutional conflicts from divergent legitimacy standards shape international firms' environmental responsibility practices, intensifying institutional pressures (Brammer, Pavelin and Porter, 2009; Tan and Wang, 2011; Ahmadova *et al.*, 2023). Maintaining legitimacy, thus, becomes challenging during internationalization, intensifying institutional pressures (Campbell, 2007; Tan and Wang, 2011). Hence, internationalizing firms must enhance stakeholder engagement, especially under high internationalization, prompting the development of environmental strategies and increased vigilance against unethical behavior in host countries (Brammer, Pavelin and Porter, 2009; Duque-Grisales *et al.*, 2020). In response, firms are motivated to utilize BDA to enhance environmental performance and meet stakeholder needs. Leveraging BDA provides deeper insights into environmental trends and stakeholder preferences, enabling more effective environmental strategies aligned with stakeholder expectations (Cheng et al., 2020; Gnizy, 2019), fostering trust and credibility among stakeholders (Saeed, Riaz and Baloch, 2022), and ultimately enhancing their overall environmental performance.

From the supply side, extensive internationalization offers firms the opportunities to leverage BDA for improving environmental performance. Initially, through international engagements, firms gain market insights and technological expertise crucial for environmental initiatives(Salomon and Jin, 2010). Firms with heightened internationalization levels across diverse countries encounter various environmental issues and sustainability practices (Campbell, 2007). This diversity grants access to a broad spectrum of environmental data, encompassing climate patterns, regulatory landscapes, and ecological challenges. BDA's data integration aspect aggregates information from various sources, enabling firms to develop sophisticated predictive models for reducing carbon footprints, optimizing resource usage, and enhancing sustainability practices (Popovič *et al.*, 2012). Firms armed with richer datasets can develop more sophisticated and accurate predictive models using BDA, leading to more effective strategies for reducing carbon footprints, optimizing resource usage, and enhancing overall sustainability practices (Wang et al., 2023). Moreover, the analytical capacity of BDA plays a vital role in enhancing decision-making related to the customization and implementation of environmental initiatives. International diversity necessitates tailoring products, services, and sustainability strategies to meet the specific demands and regulations of different markets (Husted and Allen, 2006). BDA provides insights into local preferences and environmental concerns, enabling the development of customized sustainability initiatives. Leveraging analytical capabilities, firms convert descriptive information into actionable knowledge, offering practical decision-making options (Gudfinnsson, Strand and Berndtsson, 2015). Advanced analytics predict future regulatory trends based on current data, enabling proactive alignment of environmental strategies with potential regulations to avoid compliance issues.

Hypothesis 2. *Internationalization positively moderates the impact of BDA utilization on environmental performance.*

3. Data and Methodology

3.1. Data

We compile a distinctive panel dataset encompassing Fortune 500 firms as the target firms between 2010 and 2021 due to their extensive coverage by the media (Joshi *et al.*, 2010). The selection of 2010 as the starting point is motivated by the initiation of mass media's heightened interest in big data during that year. Information regarding firm-level BDA announcements is aggregated from reputable U.S. news outlets, with Lexis Nexis and Factiva serving as newspaper databases archiving news from these sources. Air pollutants data and greenhouse gas (GHG) emissions are sourced from the S&P Capital IQ Pro Environmental Package. Financial data are sourced from Compustat, resulting in a sample of 249 firms with 2,882 firm-year observations.

3.2. Variable definitions

3.2.1. Dependent variable

We operationalize environmental performance using *external costs of air pollutants* and *GHG emissions* as proxies. External costs represent unremunerated impacts from production or consumption of a commodity upon a non-participating third party, quantified annually in US dollars by S&P (Azhgaliyeva and Le, 2023). Confirming our hypothesis would entail a negative coefficient, reflecting the anticipated positive impact of BDA on reducing external costs linked to air pollutants and GHG emissions, signaling enhanced environmental performance.

3.2.2. Explanatory variable

BDA utilization is operationally defined as the annual cumulative count of BDA announcements per firm retrieved from Lexis Nexis and Factiva databases, consistent with prior research methodologies (Chi, Ravichandran and Andrevski, 2010; Joshi *et al.*, 2010) (see more in Appendix 1).

3.2.3. Moderating variables

Internationalization degree is measured as the ratio of foreign to total sales (Abdi and Aulakh, 2018). *Internationalization speed* is determined by dividing the number of foreign subsidiaries by the duration since the company's initial expansion into foreign markets (Vermeulen and Barkema, 2002; Casillas and Moreno-Menéndez, 2014). *International scope* is measured by the count of countries in which the firm established subsidiaries during its global expansion (Vermeulen and Barkema, 2002). *Internationalization rhythm* is gauged through the kurtosis of the first derivative of the firm's foreign subsidiary count over time (Vermeulen and Barkema, 2002).

$$kurtosis = \left\{\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left(\frac{x_i-\bar{x}}{s}\right)^4\right\} - \frac{3(n-1)^2}{(n-2)(n-3)},$$

where *n* represents the number of observations, x_i is the number of foreign subsidiaries of the firm in the year *i*, $\bar{x}$ and *s* are the mean and standard deviation of the number of overseas subsidiaries of the firm in the observation period [t,t + 4], respectively.

3.2.4. Control variables

Based on related literature (Walls, Berrone and Phan, 2012; Tang and Tang, 2018), we select numerous control variables, namely *firm size*, *firm profitability*, *firm age*, *financial leverage*, *sale growth*, *capital intensity*, *R&D intensity*, and *entrepreneurial orientation* (see more in Appendix 2 and Appendix 3). A summary of descriptive Statistics is shown in Appendix 4.

3.3. Empirical strategies

We present the following econometric model to examine the impact of BDA utilization on environmental performance:

$$EP_{ijt} = \beta_0 + \beta_1\, bda_{ijt} + \boldsymbol{\beta_2}'\boldsymbol{X_{ijt}} + \lambda_i + \lambda_{jt} + \varepsilon_{i,t}\,, (1)$$

where EP_{ijt} represents a firm's environmental performance. bda_{ijt} denotes the BDA. $\boldsymbol{X_{ijt}}$ symbolizes the set of control variables. λ_i and λ_{jt} are individual fixed effect and year sector fixed effect, respectively, and $\varepsilon_{i,t}$ is an error term.

4. Results

4.1. Benchmark regression results

This section explores the effects of employing BDA on a firm's environmental performance. The findings from two distinct models are presented in Table 1.[1] The results substantiate that enhanced BDA utilization leads to improved environmental performance. For instance, regarding air pollutants, both random-effect and fixed-effect regressions show point estimates of -0.104 and -0.105, respectively. This demonstrates that a unit increase in BDA utilization leads to a 10% decrease in a firm's air pollutants, thereby validating Hypothesis 1. Similar outcomes are observed when evaluating environmental performance using GHG emissions as a proxy (see more in Appendix 5).

(insert Table 1 here)

4.2. Endogeneity concerns

While the fixed-effect regressions yield appealing results, there might be potential endogeneity concerns due to omitted variable biases and measurement errors. To address these endogeneity issues, we utilize the two-stage least-squares (2SLS) regression by using the industry median of BDA utilization as an instrumental variable (IV) for firms' BDA utilization.[2] Column 1 of Table 2 presents the estimated results. The negative impacts of BDA utilization on firms' air pollutants and GHG emissions persist, even after accounting for endogeneity concerns.

To enhance the validity of the outcomes derived from 2SLS estimates, we undertake two additional robustness checks. First, we utilize the heteroskedasticity-based instrument of Lewbel (2012) to mitigate the potential endogeneity problem. The results are reported in Columns 2 and 5 of Table 2, demonstrating a consistent and statistically significant negative nexus between BDA utilization and firms' air pollutants (or GHG emissions). Second, we utilize the lagged BDA utilization as an additional IV. Results are reported in Columns 3 and 6 of Table 2. We consistently document that BDA utilization negatively affects firms' air pollutants (or GHG emissions).

(insert Table 2 here)

4.3. Heterogenous effects

This section investigates the heterogeneous effects of BDA across key dimensions of internationalization, encompassing degree, speed, rhythm, and scope. For each dimension, we

---

[1] The Hausman specification test indicates a preference for fixed effects regression over random effects regression when assessing the impact of BDA utilization on firms' environmental performance.

[2] The utilization of industry variables as IV has been documented in prior scholarly works, as evidenced by Azhgaliyeva and Le (2023), Cai et al. (2011), and Ferrat (2021).

categorize firms into two subsamples: low and high. The heterogenous impacts of BDA on firms' air pollutants and GHG emissions are displayed in Figure 1a and Figure 1b, respectively. Notably, regardless of internationalization dimensions or environmental performance measures, the effect of BDA utilization on firms' environmental performance is more pronounced and statistically significant for firms with high levels of internationalization when juxtaposed with their low-level internationalization counterparts. Thus, our finding highlights the importance of internationalization in the relationship between BDA utilization and environmental performance, supporting Hypothesis 2.

(insert Figure 1 here)

5. Conclusions

This study uncovers a causal relationship between firms' BDA utilization and their environmental performance. It contributes to the literature in two significant ways: first, by empirically demonstrating the favorable influence of BDA utilization on environmental performance, aligning with organizational learning theory principles; and second, by highlighting the moderating effect of internationalization, drawing from stakeholder theory and knowledge management perspectives. This nuanced insight enriches our understanding of how organizations can effectively utilize BDA to address environmental challenges in an increasingly interconnected world.

To enhance BDA's role in improving environmental performance, policymakers should consider several recommendations. First, policymakers should incentivize organizations to invest in robust BDA infrastructure to bolster their capacity for data collection, analysis, and insights generation. Second, governments and industry bodies can facilitate knowledge sharing and collaboration initiatives among organizations to promote best practices in BDA utilization for environmental sustainability, while also encouraging cross-border collaboration and partnerships to leverage BDA's positive impact on environmental performance, emphasizing international research projects and joint ventures. Third, empirical results of this study underscore the urgency of addressing environmental impacts, especially among rapidly growing larger firms. This could involve fostering partnerships between small and large firms for knowledge transfer and resource sharing, alongside emphasizing prudent financial management to mitigate environmental footprint and support sustainable growth through efficient resource allocation.

This study acknowledges limitations and suggests future research directions. While our findings are based on the US context, exploring other institutional settings could provide deeper insights. Future studies could investigate variables and relationships across diverse economies, integrating an institutions-based view with dynamic capabilities. Furthermore, comparative analyses across various firm types, including SMEs, family-owned enterprises, and multinational corporations, could illuminate different approaches to value creation through big data. Qualitative methods, such as ethnography and interviews, could enhance understanding of big data management challenges and their impact on environmental performance. Moreover, the lack of statistical significance observed in some of the control variables across all models highlights a limitation in our study, suggesting an avenue for future research to explore alternative control variables or model specifications to better understand their influence on environmental performance.

Reference

3p Contributors (2016) 'How big data can make your company more sustainable', *Triple Pundit: People, Planet and Profit (Feb 10th 2016)*. Available at: https://www.triplepundit.com/story/2016/how-big-data-can-make-your-company-more-sustainable/28671.

Abdi, M. and Aulakh, P.S. (2018) 'Internationalization and performance: Degree, duration, and scale of operations', *Journal of International Business Studies*, 49(7), pp. 832–857. Available at: https://doi.org/10.1057/s41267-018-0146-8.

Aguilera-Caracuel, J., Hurtado-Torres, N.E. and Aragón-Correa, J.A. (2012) 'Does international experience help firms to be green? A knowledge-based view of how international experience and organisational learning influence proactive environmental strategies', *International Business Review*, 21(5), pp. 847–861. Available at: https://doi.org/10.1016/j.ibusrev.2011.09.009.

Ahi, A.A. *et al.* (2022) 'Advanced technologies and international business: A multidisciplinary analysis of the literature', *International Business Review*, 31(4), p. 101967. Available at: https://doi.org/10.1016/j.ibusrev.2021.101967.

Ahmadova, G. *et al.* (2023) 'Firm- and Country-Specific Advantages: Towards a Better Understanding of MNEs' Environmental Performance in the International Arena', *Organization & Environment*, 36(3), pp. 468–497. Available at: https://doi.org/10.1177/10860266221129699.

Akter, S. *et al.* (2016) 'How to improve firm performance using big data analytics capability and business strategy alignment?', *International Journal of Production Economics*, 182, pp. 113–131. Available at: https://doi.org/10.1016/j.ijpe.2016.08.018.

Ali, M. (2021) 'Imitation or innovation: To what extent do exploitative learning and exploratory learning foster imitation strategy and innovation strategy for sustained competitive advantage?☆', *Technological Forecasting and Social Change*, 165, p. 120527. Available at: https://doi.org/10.1016/j.techfore.2020.120527.

Asimakopoulos, G., Revilla, A.J. and Slavova, K. (2020) 'External Knowledge Sourcing and Firm Innovation Efficiency', *British Journal of Management*, 31(1), pp. 123–140. Available at: https://doi.org/10.1111/1467-8551.12367.

Asmussen, C.G. and Fosfuri, A. (2019) 'Orchestrating corporate social responsibility in the multinational enterprise', *Strategic Management Journal*, 40(6), pp. 894–916. Available at: https://doi.org/10.1002/smj.3007.

Azhgaliyeva, D. and Le, H. (2023) *Investment in Renewable Energy and Emissions: Firm-Level Empirical Evidence from the People's Republic of China*. Asian Development Bank Institute. Available at: https://doi.org/10.56506/INEL6435.

Bansal, P. (2005) 'Evolving sustainably: a longitudinal study of corporate sustainable development', *Strategic Management Journal*, 26(3), pp. 197–218. Available at: https://doi.org/10.1002/smj.441.

Barbosa, M.W. *et al.* (2022) 'The effects of internationalization orientation in the sustainable performance of the agri-food industry through environmental collaboration: An emerging economy perspective', *Sustainable Production and Consumption*, 31, pp. 407–418. Available at: https://doi.org/10.1016/j.spc.2022.03.013.

Belhadi, A. *et al.* (2020) 'The integrated effect of Big Data Analytics, Lean Six Sigma and Green Manufacturing on the environmental performance of manufacturing companies: The case of North Africa', *Journal of Cleaner Production*, 252, p. 119903. Available at: https://doi.org/10.1016/j.jclepro.2019.119903.

Blomstermo, A. *et al.* (2004) ‘The perceived usefulness of network experiential knowledge in the internationalizing firm’, *Journal of International Management*, 10(3), pp. 355–373. Available at: https://doi.org/10.1016/j.intman.2004.05.004.

Boso, N., Story, V.M. and Cadogan, J.W. (2013) ‘Entrepreneurial orientation, market orientation, network ties, and performance: Study of entrepreneurial firms in a developing economy’, *Journal of Business Venturing*, 28(6), pp. 708–727. Available at: https://doi.org/10.1016/j.jbusvent.2013.04.001.

Brammer, S.J., Pavelin, S. and Porter, L.A. (2009) ‘Corporate Charitable Giving, Multinational Companies and Countries of Concern’, *Journal of Management Studies*, 46(4), pp. 575–596. Available at: https://doi.org/10.1111/j.1467-6486.2008.00827.x.

Cai, Y., Jo, H. and Pan, C. (2011) ‘Vice or Virtue? The Impact of Corporate Social Responsibility on Executive Compensation’, *Journal of Business Ethics*, 104(2), pp. 159–173. Available at: https://doi.org/10.1007/s10551-011-0909-7.

Calic, G. and Ghasemaghaei, M. (2021) ‘Big data for social benefits: Innovation as a mediator of the relationship between big data and corporate social performance’, *Journal of Business Research*, 131, pp. 391–401. Available at: https://doi.org/10.1016/j.jbusres.2020.11.003.

Campbell, J.L. (2007) ‘Why would corporations behave in socially responsible ways? an institutional theory of corporate social responsibility’, *Academy of Management Review*, 32(3), pp. 946–967. Available at: https://doi.org/10.5465/amr.2007.25275684.

Casillas, J.C. and Moreno-Menéndez, A.M. (2014) ‘Speed of the internationalization process: The role of diversity and depth in experiential learning’, *Journal of International Business Studies*, 45(1), pp. 85–101. Available at: https://doi.org/10.1057/jibs.2013.29.

Cheng, C., Zhong, H. and Cao, L. (2020) ‘Facilitating speed of internationalization: The roles of business intelligence and organizational agility’, *Journal of Business Research*, 110, pp. 95–103. Available at: https://doi.org/10.1016/j.jbusres.2020.01.003.

Cheng, J. and Liu, Y. (2018) ‘The effects of public attention on the environmental performance of high-polluting firms: Based on big data from web search in China’, *Journal of Cleaner Production*, 186, pp. 335–341. Available at: https://doi.org/10.1016/j.jclepro.2018.03.146.

Chi, L., Ravichandran, T. and Andrevski, G. (2010) ‘Information Technology, Network Structure, and Competitive Action’, *Information Systems Research*, 21(3), pp. 543–570. Available at: https://doi.org/10.1287/isre.1100.0296.

Dam, N.A.K., Le Dinh, T. and Menvielle, W. (2019) ‘A systematic literature review of big data adoption in internationalization’, *Journal of Marketing Analytics*, 7(3), pp. 182–195. Available at: https://doi.org/10.1057/s41270-019-00054-7.

Dubey, R. *et al.* (2019) ‘Can big data and predictive analytics improve social and environmental sustainability?’, *Technological Forecasting and Social Change*, 144, pp. 534–545. Available at: https://doi.org/10.1016/j.techfore.2017.06.020.

Dubey, R. *et al.* (2020) ‘Big data analytics and artificial intelligence pathway to operational performance under the effects of entrepreneurial orientation and environmental dynamism: A study of manufacturing organisations’, *International Journal of Production Economics*, 226, p. 107599. Available at: https://doi.org/10.1016/j.ijpe.2019.107599.

Duque-Grisales, E. *et al.* (2020) ‘Can proactive environmental strategy improve Multilatinas’ level of internationalization? The moderating role of board independence’, *Business Strategy and the Environment*, 29(1), pp. 291–305. Available at: https://doi.org/10.1002/bse.2377.

Emani, C.K., Cullot, N. and Nicolle, C. (2015) ‘Understandable Big Data: A survey’, *Computer Science Review*, 17, pp. 70–81. Available at: https://doi.org/10.1016/j.cosrev.2015.05.002.

Ferrat, Y. (2021) ‘Carbon emissions and firm performance: A matter of horizon, materiality and regional specificities’, *Journal of Cleaner Production*, 329, p. 129743. Available at: https://doi.org/10.1016/j.jclepro.2021.129743.

Freeman, E. and Liedtka, J. (1997) ‘Stakeholder capitalism and the value chain’, *European Management Journal*, 15(3), pp. 286–296. Available at: https://doi.org/10.1016/S0263-2373(97)00008-X.

George, G. and Schillebeeckx, S.J.D. (2022) ‘Digital transformation, sustainability, and purpose in the multinational enterprise’, *Journal of World Business*, 57(3), p. 101326. Available at: https://doi.org/10.1016/j.jwb.2022.101326.

Gnizy, I. (2019a) ‘Big data and its strategic path to value in international firms’, *International Marketing Review*, 36(3), pp. 318–341. Available at: https://doi.org/10.1108/IMR-09-2018-0249.

Gnizy, I. (2019b) ‘Big data and its strategic path to value in international firms’, *International Marketing Review*, 36(3), pp. 318–341. Available at: https://doi.org/10.1108/IMR-09-2018-0249.

González-Benito, J. and González-Benito, Ó. (2006) ‘A review of determinant factors of environmental proactivity’, *Business Strategy and the Environment*, 15(2), pp. 87–102. Available at: https://doi.org/10.1002/bse.450.

Gudfinnsson, K., Strand, M. and Berndtsson, M. (2015) ‘Analyzing business intelligence maturity’, *Journal of Decision Systems*, 24(1), pp. 37–54. Available at: https://doi.org/10.1080/12460125.2015.994287.

Gupta, S. *et al.* (2019) ‘Circular economy and big data analytics: A stakeholder perspective’, *Technological Forecasting and Social Change*, 144, pp. 466–474. Available at: https://doi.org/10.1016/j.techfore.2018.06.030.

Hernández, Á.B. *et al.* (2018) ‘Using machine learning to optimize parallelism in big data applications’, *Future Generation Computer Systems*, 86, pp. 1076–1092. Available at: https://doi.org/10.1016/j.future.2017.07.003.

Huber, G.P. (1991) ‘Organizational Learning: The Contributing Processes and the Literatures’, *Organization Science*, 2(1), pp. 88–115. Available at: https://doi.org/10.1287/orsc.2.1.88.

Husted, B.W. and Allen, D.B. (2006) ‘Corporate social responsibility in the multinational enterprise: strategic and institutional approaches’, *Journal of International Business Studies*, 37(6), pp. 838–849. Available at: https://doi.org/10.1057/palgrave.jibs.8400227.

Joshi, K.D. *et al.* (2010) ‘Changing the Competitive Landscape: Continuous Innovation Through IT-Enabled Knowledge Capabilities’, *Information Systems Research*, 21(3), pp. 472–495. Available at: https://doi.org/10.1287/isre.1100.0298.

Kogut, B. (1989) ‘Research notes and communications a note on global strategies’, *Strategic Management Journal*, 10(4), pp. 383–389. Available at: https://doi.org/10.1002/smj.4250100407.

Koseleva, N. and Ropaite, G. (2017) ‘Big Data in Building Energy Efficiency: Understanding of Big Data and Main Challenges’, *Procedia Engineering*, 172, pp. 544–549. Available at: https://doi.org/10.1016/j.proeng.2017.02.064.

Kristoffersen, E. *et al.* (2021) ‘The effects of business analytics capability on circular economy implementation, resource orchestration capability, and firm performance’, *International Journal of Production Economics*, 239, p. 108205. Available at: https://doi.org/10.1016/j.ijpe.2021.108205.

Lewbel, A. (2012) ‘Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models’, *Journal of Business & Economic Statistics*, 30(1), pp. 67–80. Available at: https://doi.org/10.1080/07350015.2012.643126.

Lubinski, C. and Wadhwani, R.D. (2020) ‘Geopolitical jockeying: Economic nationalism and multinational strategy in historical perspective’, *Strategic Management Journal*, 41(3), pp. 400–421. Available at: https://doi.org/10.1002/smj.3022.

Mikalef, P. *et al.* (2019) ‘Big data analytics and firm performance: Findings from a mixed-method approach’, *Journal of Business Research*, 98, pp. 261–276. Available at: https://doi.org/10.1016/j.jbusres.2019.01.044.

Ocelík, V., Kolk, A. and Ciulli, F. (2023) ‘Multinational enterprises, Industry 4.0 and sustainability: A multidisciplinary review and research agenda’, *Journal of Cleaner Production*, 413, p. 137434. Available at: https://doi.org/10.1016/j.jclepro.2023.137434.

Phillips-Wren, G. and Hoskisson, A. (2015) ‘An analytical journey towards big data’, *Journal of Decision Systems*, 24(1), pp. 87–102. Available at: https://doi.org/10.1080/12460125.2015.994333.

Popovič, A. *et al.* (2012) ‘Towards business intelligence systems success: Effects of maturity and culture on analytical decision making’, *Decision Support Systems*, 54(1), pp. 729–739. Available at: https://doi.org/10.1016/j.dss.2012.08.017.

Queiroz, M.M. (2018) ‘A framework based on Twitter and big data analytics to enhance sustainability performance’, *Environmental Quality Management*, 28(1), pp. 95–100. Available at: https://doi.org/10.1002/tqem.21576.

Riggs, R. *et al.* (2023) ‘Opening the black box of big data sustainable value creation: the mediating role of supply chain management capabilities and circular economy practices’, *International Journal of Physical Distribution & Logistics Management*, 53(7/8), pp. 762–788. Available at: https://doi.org/10.1108/IJPDLM-03-2022-0098.

Roman Pais Seles, B.M. *et al.* (2018) ‘Business opportunities and challenges as the two sides of the climate change: Corporate responses and potential implications for big data management towards a low carbon society’, *Journal of Cleaner Production*, 189, pp. 763–774. Available at: https://doi.org/10.1016/j.jclepro.2018.04.113.

Saeed, A., Riaz, H. and Baloch, M.S. (2022) ‘Does big data utilization improve firm legitimacy?’, *Technological Forecasting and Social Change*, 182, p. 121847. Available at: https://doi.org/10.1016/j.techfore.2022.121847.

Salomon, R. and Jin, B. (2010) ‘Do leading or lagging firms learn more from exporting?’, *Strategic Management Journal*, 31(10), pp. 1088–1113. Available at: https://doi.org/10.1002/smj.850.

Sarker, M.N.I. *et al.* (2020) ‘Disaster resilience through big data: Way to environmental sustainability’, *International Journal of Disaster Risk Reduction*, 51, p. 101769. Available at: https://doi.org/10.1016/j.ijdrr.2020.101769.

Sivarajah, U. *et al.* (2017) ‘Critical analysis of Big Data challenges and analytical methods’, *Journal of Business Research*, 70, pp. 263–286. Available at: https://doi.org/10.1016/j.jbusres.2016.08.001.

Song, M. *et al.* (2017) ‘How would big data support societal development and environmental sustainability? Insights and practices’, *Journal of Cleaner Production*, 142, pp. 489–500. Available at: https://doi.org/10.1016/j.jclepro.2016.10.091.

Song, M., Du, Q. and Zhu, Q. (2017) ‘A theoretical method of environmental performance evaluation in the context of big data’, *Production Planning & Control*, 28(11–12), pp. 976–984. Available at: https://doi.org/10.1080/09537287.2017.1336801.

Song, M.-L. *et al.* (2018) ‘Environmental performance evaluation with big data: theories and methods’, *Annals of Operations Research*, 270(1–2), pp. 459–472. Available at: https://doi.org/10.1007/s10479-016-2158-8.

Tan, J. and Wang, L. (2011) ‘MNC Strategic Responses to Ethical Pressure: An Institutional Logic Perspective’, *Journal of Business Ethics*, 98(3), pp. 373–390. Available at: https://doi.org/10.1007/s10551-010-0553-7.

Tang, Z. and Tang, J. (2018) ‘Stakeholder Corporate Social Responsibility Orientation Congruence, Entrepreneurial Orientation and Environmental Performance of Chinese Small and Medium-sized Enterprises’, *British Journal of Management*, 29(4), pp. 634–651. Available at: https://doi.org/10.1111/1467-8551.12255.

Vermeulen, F. and Barkema, H. (2002) ‘Pace, rhythm, and scope: process dependence in building a profitable multinational corporation’, *Strategic Management Journal*, 23(7), pp. 637–653. Available at: https://doi.org/10.1002/smj.243.

Walls, J.L., Berrone, P. and Phan, P.H. (2012) ‘Corporate governance and environmental performance: is there really a link?’, *Strategic Management Journal*, 33(8), pp. 885–913. Available at: https://doi.org/10.1002/smj.1952.

Wang, H. *et al.* (2023) ‘How does digital technology promote carbon emission reduction? Empirical evidence based on e-commerce pilot city policy in China’, *Journal of Environmental Management*, 325, p. 116524. Available at: https://doi.org/10.1016/j.jenvman.2022.116524.

Wang, Y., Kung, L. and Byrd, T.A. (2018) ‘Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations’, *Technological Forecasting and Social Change*, 126, pp. 3–13. Available at: https://doi.org/10.1016/j.techfore.2015.12.019.

Wei, X., Li, H. and Wang, Y. (2024) ‘Does big data infrastructure development promote new enterprises entry: Evidence from China’, *Finance Research Letters*, 60, p. 104828. Available at: https://doi.org/10.1016/j.frl.2023.104828.

Wu, K.-J. *et al.* (2017) ‘Toward sustainability: using big data to explore the decisive attributes of supply chain risks and uncertainties’, *Journal of Cleaner Production*, 142, pp. 663–676. Available at: https://doi.org/10.1016/j.jclepro.2016.04.040.

Zhu, S., Dong, T. and Luo, X. (Robert) (2021) ‘A longitudinal study of the actual value of big data and analytics: The role of industry environment’, *International Journal of Information Management*, 60, p. 102389. Available at: https://doi.org/10.1016/j.ijinfomgt.2021.102389.

Table 1. Estimated impacts of BDA utilization on firms' environmental performance

	Air pollutants		GHG emissions	
	Random effects	Fixed effects	Random effects	Fixed effects
BDA utilization	-0.104***	-0.105***	-0.060**	-0.062**
	(0.031)	(0.031)	(0.027)	(0.027)
Firm size	0.359***	0.357***	0.341***	0.340***
	(0.051)	(0.056)	(0.051)	(0.054)
Firm profitability	0.104	0.050	0.118	0.038
	(0.278)	(0.282)	(0.273)	(0.280)
Firm age	-0.080	-0.069	0.063	0.114
	(0.121)	(0.165)	(0.094)	(0.122)
Financial leverage	-0.054	-0.101	-0.281	-0.369*
	(0.263)	(0.280)	(0.181)	(0.189)
Sale growth	0.016	0.006	0.117**	0.110*
	(0.065)	(0.064)	(0.058)	(0.059)
Capital intensity	-0.879	-1.518**	-0.011	-0.487
	(0.706)	(0.762)	(0.366)	(0.374)
R&D intensity	-0.342	0.135	-1.215**	-0.837
	(0.547)	(0.658)	(0.479)	(0.542)
Entrepreneurial orientation	-0.123	0.072	0.322	0.466
	(0.319)	(0.317)	(0.380)	(0.384)
Constant	-3.895***	-3.881***	-3.407***	-3.579***
	(0.888)	(1.075)	(0.983)	(1.121)
Hausman test	Chi2(09) = 1,067.82***		Chi2(09) = 141.76***	
Observations	2,882	2,882	2,882	2,882
R-squared		0.065		0.071
Number of companies	249	249	249	249

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. We utilize the "xttest3" and "xtserial" commands in Stata to assess heteroskedasticity and autocorrelation, respectively, and our findings indicate its presence within our sample. Consequently, we employ robust standard errors to address these issues. The Hausman specification test is utilized to examine the suitability of the fixed-effect regression in comparison to the random-effect regression.

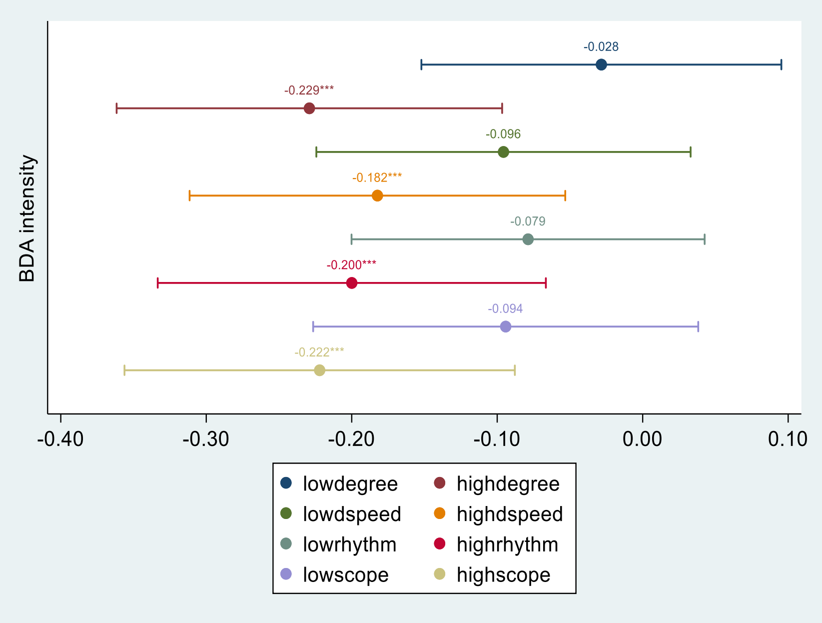
Table 2. Estimated impacts of BDA utilization on firms' environmental performance using 2SLS approach

	Air pollutants			GHG emissions		
	Industry median (1)	Industry median & internal IV (2)	Industry median & lagged BDA utilization (3)	Industry median (4)	Industry median & internal IV (5)	Industry median & lagged BDA utilization (6)
BDA utilization	-0.133***	-0.121***	-0.061**	-0.142***	-0.126***	-0.069**
	(0.047)	(0.046)	(0.030)	(0.043)	(0.041)	(0.027)
Firm size	0.384***	0.383***	0.369***	0.365***	0.364***	0.338***
	(0.030)	(0.030)	(0.033)	(0.030)	(0.030)	(0.033)
Firm profitability	0.093	0.094	-0.016	0.066	0.067	-0.042
	(0.257)	(0.257)	(0.271)	(0.224)	(0.224)	(0.212)
Firm age	-0.010	-0.015	-0.132	0.184***	0.177***	0.164*

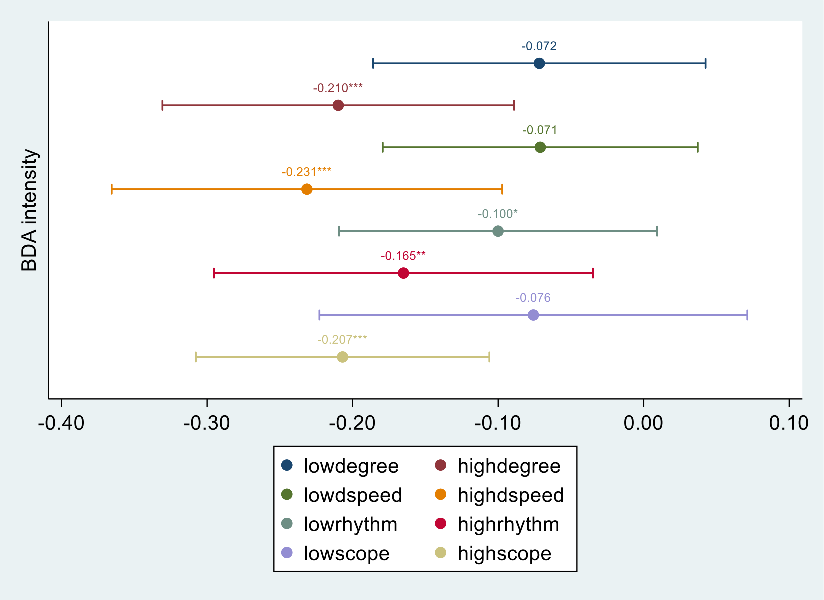
	(0.076)	(0.076)	(0.128)	(0.061)	(0.060)	(0.099)
Financial leverage	-0.036	-0.048	-0.121	-0.260**	-0.277**	-0.292**
	(0.138)	(0.139)	(0.140)	(0.122)	(0.122)	(0.119)
Sale growth	-0.004	-0.003	-0.006	0.100	0.101	0.100
	(0.060)	(0.059)	(0.064)	(0.068)	(0.068)	(0.072)
Capital intensity	-1.494***	-1.503***	-1.718***	-0.422	-0.435	-0.490
	(0.556)	(0.558)	(0.596)	(0.452)	(0.454)	(0.474)
R&D intensity	0.268	0.256	0.234	-0.678	-0.694	-0.981*
	(0.568)	(0.566)	(0.594)	(0.527)	(0.525)	(0.557)
Entrepreneurial orientation	0.147	0.136	0.072	0.577**	0.562**	0.514**
	(0.249)	(0.248)	(0.256)	(0.228)	(0.225)	(0.229)
Observations	2,882	2,882	2,664	2,882	2,882	2,664
First stage						
Industry median	0.695***	0.659***	0.199***	0.695***	0.659***	0.199***
	(0.030)	(0.029)	(0.024)	(0.030)	(0.029)	(0.024)
Lagged BDA utilization	–	–	0.703***	–	–	0.703***
	–	–	(0.019)	–	–	(0.019)
Internal IVs	–	Yes	–	–	Yes	–
Under id.	312.286	318.548	460.043	312.286	318.548	460.043
Weak id.	530.876	72.525	1081.620	530.876	72.525	1081.620
Hansen J stat	–	5.094	0.784	–	4.924	6.721

Notes: Weak id. and Under id. are weak identification test and under identification test, respectively. We report Kleibergen-Paap rk Wald F-statistics and Kleiberg-Paap rk LM statistic for Weak id. and Under id., respectively. Robust standard errors in parentheses. Column (2) presents the estimated results of the Lewbel 2SLS approach. Lagged BDA utilization refers to an additional IV. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1. Heterogeneous effects



a. Air pollutants



b. GHG emissions

Citation on deposit: Le, H., & Vu, K.-C. (2024). Big data analytics and environmental performance: The moderating role of internationalization. Finance Research Letters, 64, Article 105484. https://doi.org/10.1016/j.frl.2024.105484