PUBLISHED BY
VOPROSY
EKONOMIKI
N. P.

Russian Journal of Economics 10 (2024) 84–102
DOI 10.32609/j.ruje.10.107500
Publication date: 29 March 2024

Russian
Journal of
Economics

www.rujec.org

Examining the impact of national open data initiatives on human development: A comparative study between Latin America and Africa

Tosin Ekundayo[a,b,*], Amiya Bhaumik[b],
Justine Chinoperekweyi[c], Zafarullah Khan[a]

[a] *Synergy University Dubai, Dubai, United Arab Emirates*
[b] *Lincoln University College, Petaling Jaya, Malaysia*
[c] *Centre for Organization Leadership and Development, Maputo, Zimbabwe*

Abstract

In an era where data-driven decision-making is crucial for sustainable development, the role of open data initiatives in shaping potential and strategic outcomes has gained increasing attention. This study investigates the potential impact of National Open Data (NOD) initiatives on human capital development, with specific emphasis on their contribution towards achieving United Nations Sustainable Development Goal 3 (SDG3) targets. It explores the relationship between these initiatives and the Human Development Index (HDI) across different countries and regions aiming to ascertain if there is a significant association between open data and human development. The results indicate a strong positive correlation between NOD initiatives and HDI, suggesting that open data can play a crucial role in enhancing human development and meeting SDG3 targets. However, the strength of this relationship varies significantly across regions, with a more pronounced impact observed in Latin America compared to Africa. These findings underscore the potential of open data in propelling human capital development but also highlight the need to contextualize such initiatives based on unique regional dynamics. The study serves as a resource for policymakers in leveraging open data to enhance human development outcomes and progress towards achieving SDGs.

Keywords: open data, human development index, HDI, Sustainable Development Goal 3, SDG3.
JEL classification: F, O.

1. Introduction

The United Nations Sustainable Development Goals (SDGs) act as a global blueprint that countries across the globe can utilize to address a diverse range

* Corresponding author, E-mail address: tosinekundayo@outlook.com

of socio-economic and environmental challenges. SDG3, one of the 17 goals, is specifically designed to "guarantee a healthy lifestyle and foster well-being for all individuals across all ages." The worldwide community is tirelessly working towards realizing the SDGs, including SDG3 (Meurs et al., 2019). To accomplish such a challenging goal, it necessitates not just well-founded policy choices and efficient allocation of resources, but also the accessibility and application of premium quality data. Given the escalating evidence of the impact of open data on social and economic outcomes, decision-makers are shifting their focus towards the capacity of open data to contribute towards achieving SDG3 objectives. This shift in focus has resulted in a push for open data (Ekundayo, 2021; Ekundayo et al., 2023a). As it stands, empirical evidence exists, provided by Latifah et al. (2022), suggesting a correlation between open data and human capital development, but the exploration or quantification of this relationship remains limited.

Human development, gauged by the Human Development Index (HDI), serves as a crucial contributor to accomplishing SDG3 objectives. The concept of human capital development revolves around the enhancement and honing of human skills, knowledge, capabilities, and personal traits through education, training, and various forms of personal development. This concept underpins an investment in individuals with the objective of boosting their productive and innovative potential (Njoku et al., 2017). Economically, human capital is considered a crucial element of production, contributing to a nation's economic growth and development by bolstering the workforce's ability to execute more complex tasks, apply critical thinking, and innovate (Školudová, 2016). Investments in human capital development are often viewed as indispensable for achieving individual and societal objectives (Fan et al., 2022). As such, governments frequently deploy a range of programs and policies aimed at human capital development in order to stimulate economic growth, alleviate poverty, and foster social equality. Consequently, comprehending the relationship between open data initiatives and human capital across varying regions becomes essential.

In the past ten years, numerous governments have inaugurated National Open Data (NOD) initiatives, convinced that such endeavors could heighten transparency, endorse social inclusivity, and bolster governmental efficiency (Ekundayo et al., 2023b). Open data harbors massive transformative potential (Ekundayo, 2021; Ekundayo et al., 2023a). It carries a critical function in buttressing evidence-based decision-making, nurturing innovation, and advocating for social inclusion. Additionally, it's perceived as a tool contributing to the achievement of various SDG3 targets.

Martin et al. (2017) argue that open data can augment the availability and caliber of health-related information, thereby enabling governments, researchers, and other stakeholders to monitor progress towards health-related goals more efficiently. Huston et al. (2019) suggest that open data regarding disease prevalence, vaccination coverage, and access to healthcare can facilitate the creation of targeted interventions and allow for the surveillance of their impact over time. Such data can also heighten the efficiency and efficacy of healthcare systems by supporting data-driven decision-making and resource allocation.

According to Magalhães et al. (2017), open data on healthcare facilities, medical equipment, and the distribution of healthcare workforce can aid in identifying gaps and disparities in health service provision, thereby guiding policymakers in the equitable and efficient allocation of resources. Open data can also foster innovation in the health sector by forming a basis for the development of new tools, technologies, and methods to address health challenges.

Moreover, it can enhance public trust and accountability in the health sector by fostering transparency in decision-making and resource allocation. Making health-related data publicly accessible aids governments in showcasing their commitment to addressing health challenges and equips citizens with the means to hold them accountable for their actions (Park et al., 2017).

Nonetheless, prior studies have mainly examined the relationship between open data and human development via broad measures such as the HDI or equivalent metrics, omitting the need for targeted research explicitly scrutinizing the role of open data in the accomplishment of SDG3 targets through a human capital development. A significant portion of the literature on open data and human development concentrates on single countries or regions, thereby constraining the general applicability of the findings. This limitation underscores the necessity for research that employs a more global perspective, integrating data from diverse regions to paint a more comprehensive picture of the interplay between open data initiatives and human capital development.

This study aims to enrich the existing literature by comparing the impact of open data on human capital development in Latin America and Africa. These regions are considered suitable for the comparative study due to the availability of robust data for an experimental study. Focusing on SDG3 targets, it seeks to identify factors that make open data initiatives effective in improving health outcomes. The findings will offer actionable insights for policymakers to better leverage open data for human development.

Accordingly, this paper endeavors to make a valuable contribution to the extant literature on open data and human development. It proposes a comparative analysis of the impact of open data on human capital development across diverse regions, with an aim to gain meaningful insights into the factors influencing the efficacy of open data initiatives in propelling health outcomes and achieving SDG3 targets (Ojo et al., 2022). By addressing these conspicuous research gaps, this study intends to provide a holistic understanding of the influence exerted by open data on human capital development, with a distinctive focus on realizing the SDG3 targets. More specifically, a principal objective of the study involves a comparative investigation of the relationship between NOD initiatives and the HDI in two distinctly characterized regions, Latin America and Africa. On the basis of the study's findings, recommendations can be formulated, offering vital insights that would serve as practical guides for policymakers and other relevant stakeholders aspiring to leverage the potency of open data to enhance human development in accordance with SDG3.

This study is important for its theoretical and empirical contributions, offering a model for actionable policy guidance. Its findings extend beyond the regions studied, providing globally relevant insights for achieving SDGs, particularly in improving health and human development.

This paper seeks to proffer answers as follows: Is there a significant relationship between the level of NOD initiatives and the Human Development in Latin America and Africa? Do differences in NOD initiatives in Latin America and Africa explain variations in their progress towards achieving SDG3 targets?

The structure of this study is organized to facilitate a coherent and in-depth exploration of the impact of NOD initiatives on human capital development in alignment with SDG3 targets. Section 2 begins with an integrated literature review, where existing research is discussed and further discusses theoretical and conceptual framework of the research; Section 3 is Methodology, outlining the research design based on Saunders et al. (2007) research onion model; Data analysis and model construction, detailing the regression analysis model used; Section 4 is Results, presenting the findings; Section 5 is Discussion and conclusion, interpreting those findings in the context of Latin America and Africa. Finally, Recommendations Section 6 provides actionable insights based on the study's findings, aimed at policymakers and stakeholders.

2. Literature review and research framework

The literature on open data and human development is rare; however, several studies have explored the relationship between open data and human development at different levels, including global, regional, and country-specific analyses.

Husein et al. (2015) posit that the Indonesian government needs to provide more accessible open data to small and medium enterprises (SMEs), especially for marketing information, to help them become more innovative and quality-oriented in their businesses. The research highlights the importance of open data for SMEs to improve transparency and increase public participation and also enhance productivity and competitiveness of SMEs in Indonesia, particularly for fishery SMEs. To support the argument, the paper does not explicitly mention a single theory used in this study. However, it describes the development strategy of SMEs in the Indonesian fishery sector, which can be achieved using the combination of SWOT strategies. Alamsyah et al. (2018) support this argument but focuses on HDI value prediction and its clustered nature in Indonesia.

Virkar et al. (2018) conducted a systematic literature review of 60 articles to explore the impacts of open government data on different domains, including social capital, economic development, and good governance. The study cites citizen empowerment and democracy as core expected effects of open government data (OGD) initiatives and also cited economic development as the second most commonly studied impact (Virkar et al., 2018). The authors note that the literature lacks substantial evidence of open data's impact on human development, especially in the context of achieving SDG3 targets.

Kamil et al. (2020) investigate the relationship between government openness, digitalization, and the HDI in ASEAN countries using panel data regression analysis. The findings reveal a significant positive relationship between government openness, digitalization, and human development in the these countries. The research highlights the potential benefits of pursuing policies that promote openness and digitalization as a means of achieving higher levels of human development. To support the argument, the study cites the theory of economic openness which is a concept in economics that deals with the degree to which a country or an economy engages in international trade and financial transactions. A key limitation of the study is that it focuses only on ASEAN countries, therefore the findings may not be generalizable to other regions. The study also relies on panel data rather than primary data, which could limit the accuracy and validity of the findings.

Shabbir et al. (2020) explores the scope and status of open data system in Pakistan after Right of Information Act Bill (2016), especially its impact during the recent COVID-19 pandemic and determine the role of open data system in achieving and solving social issues of Pakistan, specifically those related to transparency, accountability and good governance with a special focus on the issues generated by the COVID-19. The research concludes that open data systems can contribute to achieving transparency, accountability, and good governance, as well as fostering innovation and creativity and enhancing social development. This study did not apply any theory but used the UK outcome as the theoretical foundation of its argument as well as its base of argument. However, the paper does not extensively discuss the specific social issues and challenges that Pakistan is facing and how open data can address them.

Repkova Stofkova and Stofkova (2020) explore the use of open data in the development of the digital economy in the knowledge society in Slovak Republic. The study examines how new information communication technologies can improve the quality of life for citizens and businesses, and how open data can be used to achieve these goals. Without the application of known theory as a base of argument, the findings posit the Slovak Republic is making progress in the right direction and open data are data available to everyone, and can be used, reused and distributed without restriction.

This growing body of literature on open data and human development in pursuance of SDG3 objectives, highlights the potential of open data to contribute to achieving SDG3 targets.

To lay the theoretical foundation of this study, we begin by drawing from the Open Data for Development (OD4D) theory proposed by Sebubi et al. (2020). The OD4D theory posits that open data—defined as data that can be freely used, modified, and shared by anyone for any purpose—can significantly contribute to social, economic, and environmental development (Criado et al., 2021). This theory is instrumental in forming the basis of our inquiry into the relationship between NOD initiatives and human capital development, specifically in relation to achieving SDG3. The OD4D theory highlights that the accessibility, availability, and applicability of data can foster innovation, transparency, and collaboration (Park et al., 2017; Saxena et al., 2018). This framework acknowledges that the impacts of open data are likely influenced by a complex interplay of factors beyond the data itself, including social, political, economic, and technological contexts (de Beer, 2017). However, it guides the study in its examination of the impacts of open data on human development, taking a holistic perspective that is in line with the principles of the OD4D theory. This theoretical perspective supports the study's hypothesis that nations that proactively engage in open data initiatives might generate environments conducive to human capital development.

A conceptual framework is an illustrative device that helps to clarify the concepts that will be examined in a study, as well as the relationships among these concepts. In this study, we are investigating the impact of open data initiatives on human capital development within the context of achieving United Nations' Sustainable Development Goal 3 (SDG3). Based on the OD4D theory, the conceptual framework for this study could be represented as follows:

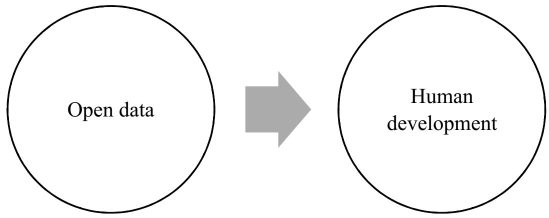

Fig. 1. Study's conceptual framework.

Source: Compiled by the authors.

Open Data Initiatives (independent variable): This component of the conceptual framework represents national actions and commitments towards implementing and maintaining open data. This is measured through data availability, accessibility, and applicability, as per the Global Open Data Index (GODI).

Human Capital Development (dependent variable): This component is the outcome that the study seeks to understand in relation to open data initiatives. Human capital development is represented through key dimensions like health, education, and income levels, as measured by the HDI.

As illustrated in Fig. 1, the conceptual framework posits that open data initiatives (IV) influence human capital development (DV) within the context of SDG3 targets, and this relationship is possibly mediated and moderated by various other factors.

The hypotheses are:

H1: There is a positive and significant relationship between the level of NOD initiatives and the human development.

H2: Differences in NOD initiatives and the human development nexus in Latin America and Africa significantly explain variations in their progress towards achieving SDG3 targets.

3. Methodology

3.1. Research design

Adopting Saunders et al. (2007) research onion model, this study carries a positivist philosophical approach to examine the correlation between open data initiatives and human capital development in alignment with SDG3 targets. It utilizes a deductive methodology, where pre-existing theories and literature form the basis for hypothesis development. These hypotheses are then tested through secondary data analysis. The study is quantitative, employing a single method, regression analysis, to evaluate the impact of open data initiatives on human development indicators. This mono-method approach enables a detailed exploration across diverse countries and regions. As a cross-sectional study, it offers a snapshot of the relationship between open data initiatives and human capital development at a specific point in time. The data collection leverages secondary data from the Global Data Barometer Index, the global evaluator of open data practices, and the UN's HDI for the year 2021. Regression analysis is used for data interpretation and hypothesis testing, aiming to present a comprehensive view of open data's influence on human development in pursuit of SDG3 targets.

Table 1
Description of variables.

Indicator	Description	Unit
NOD	Measure the state of open data initiative on a national scale	1–100
HDI	Measures average achievement in key dimensions of human development	1–100

Source: Compiled by the authors.

3.2. Data analysis and model construction

The study implements a two-stage model to investigate the impact of NOD, as per the GODI, on the HDI sourced from the United Nations Development Programme (UNDP). At the first stage, a quantitative analysis determines the correlation between open data levels and overall human development within countries. The most recent available data from both GODI and HDI are utilized to assure consistency and accuracy in the comparative analysis. This examination of the effect of NOD on HDI contributes to achieving SDG3, underscoring the potential of open data initiatives to foster human development. At the second stage, the dataset is processed to ensure suitability for analysis by confirming the availability of corresponding data for both NOD and HDI in different countries. This leads to a final sample size of 101 countries. For the Latin America and Africa comparative analysis, the variable for NOD and HDI amounts to 21 and 23 in each region. This method follows the precedents set by Constantine (2012), Ding (2006), Hess and Hess (2017), Sedgwick (2013). The variables used in this study are NOD and the HDI.

More details on description of these variables are included in Table 1.

In this study, the third phase involves conducting a correlation analysis between NOD and the HDI to verify the existence of any relationship. Subsequently, the fourth phase employs a statistical regression analysis to establish the causal relationship between NOD and HDI across 101 countries, using the most recent data available. The Ordinary Least Squares (OLS) model is used to understand the potential impact.

Following the methods of Ding (2006), Turóczy and Marian (2012), Mustapha et al. (2019), the regression model used is as follows: $Y = a + bX$. Given the dataset's size (refer to Appendix A), Microsoft Excel is the tool chosen for carrying out correlation and regression analyses for this study.

4. Results and interpretations

This section presents the results and interpretation of the statistical analyses carried out to investigate the impact of NOD initiatives on the HDI in different regions. Specifically, it includes a comparative study of Latin America and Africa, given their contrasting contexts regarding open data adoption and human development trends. Consequently, the results of these analyses are essential for testing the hypotheses and addressing the research questions of this study.

H1: *There is a positive and significant relationship between the level of NOD initiatives and the Human Development* (dataset in Appendix A).

The NOD and the HDI present differing descriptive statistical profiles on global scale (see Table 2). The average NOD score across the sample of countries is 34.09, much higher than the HDI's mean value of 0.75, illustrating a disparate scale between these two metrics. The median values for both variables — 33.32 for NOD and 0.76 for HDI — further underscore this point, reflecting different distributions for each set of scores. Examining the dispersion of the data, the standard deviation for NOD is considerably higher (16.03) than that for HDI (0.14), suggesting a greater spread of NOD scores around the mean compared to the HDI scores. Similarly, the sample variance for NOD is larger, revealing more variability among the NOD scores.

Both the NOD and HDI distributions demonstrate negative kurtosis, with NOD at –0.92 and HDI at –0.69, which suggests fewer outliers in both data sets than would be found in a normal distribution. Regarding skewness, the NOD data exhibits a slight right skew (0.32), whereas the HDI data is slightly left-skewed (–0.42), indicating differing asymmetries in their distributions. The range of scores for both variables — from 6.41 to 68.02 for NOD, and from 0.45 to 0.95 for HDI — highlights the diverse circumstances and policies of the countries under study.

These comparative insights are crucial for this study as they not only provide a comprehensive understanding of the central tendencies and dispersions for NOD and HDI, but also set the stage for regression analysis by revealing the underlying data distributions.

Globally, the correlation coefficient of 0.705829 between the NOD and HDI signifies a strong positive relationship between these two variables (see Table 3). This means that countries with higher NOD scores, indicating a more robust open data policy, tend to also have higher HDI scores, which represent better human development outcomes. The correlation value lies between 0 and 1, where 0 suggests no correlation and 1 implies a perfect positive correlation, so the result of 0.705829 is quite significant.

The implication of this correlation for the study is important. It indicates the potential of open data initiatives as a tool for improving human development outcomes. However, while this strong correlation is noteworthy, it does not establish a causal relationship. The regression analysis is required for further investigation of this possible cause-effect relationship. The positive correlation does suggest that the regression analysis may indeed find that increased NOD is associated with increased HDI, but further investigation is needed.

Table 2
Descriptive statistics for H1 on global analysis.

Indicator	Mean	Std. error	Median	Std. dev.	Sample variance	Kurtosis	Skewness	Min	Max
NOD	34.09	1.59	33.32	16.03	257.08	–0.92	0.32	6.41	68.02
HDI	0.75	0.01	0.76	0.14	0.02	–0.69	–0.42	0.45	0.95

Source: Authors' calculations.

Table 3
Pearson correlation for H1.

Variable	NOD	HDI
NOD	1	
HDI	0.705829	1

Source: Authors' calculations.

Table 4
OLS Regression for H1.

R-squared	Adjusted R-squared	F-statistics	Probability (F-statistics)	P-value
0.71	0.49	98.28	0.00	0.00

Variable	β	Std. error	t-statistics	Probability
NOD	0.01	0.00	9.91	0.00

Source: Authors' calculations.

The regression analysis output presents compelling evidence for a significant correlation between the NOD initiative and HDI scores (Table 4). The coefficient of determination (R-squared) stands at 0.71, indicating that about 71% of the variability in HDI can be explained by NOD — a strong fit for the model. Although the adjusted R-squared (accounting for the number of predictors) is somewhat lower at 49%, it remains a significant figure.

Furthermore, the F-statistics value of 98.28, coupled with a practically zero probability, rejects the null hypothesis that all regression coefficients are zero. This provides compelling evidence of a statistically significant relationship between NOD and HDI. However, the P-value is 0.00, suggesting that the model is statistically significant. The beta coefficient, meanwhile, is 0.01, signifying that for each unit increase in NOD, an expected increase of 0.01 in HDI is projected, assuming all other factors remain constant. This prediction is reinforced by the t-statistics value of 9.91 and a probability of 0.00, confirming that NOD is indeed a significant predictor of HDI.

Taken together, these regression results strongly affirm the hypothesis that open data initiatives (NOD) positively impact human development outcomes (HDI). This finding substantiates the research's primary proposition, reinforcing the central role of open data initiatives in driving human development.

One of the key assumptions in OLS regression is that the errors (or residuals) are not correlated across observations, also known as the assumption of no autocorrelation (Turner, 2020). Violation of this assumption can lead to inefficient parameter estimates. The Durbin–Watson test yielded a value of 2.071928278 (see Appendix A).

H2: *Differences in NOD initiatives and the Human development nexus in Latin America and Africa significantly explain variations in their progress towards achieving SDG3 targets* (dataset in Appendix B and C)

The descriptive statistics for the two regions, Latin America and Africa, provide a comparative analysis of their NOD initiatives and HDI values, relative to their progress towards SDG3 targets. In Latin America, the mean NOD value is 32.71, which is significantly higher than Africa's mean of 19.97. This indicates that Latin American countries, on average, have more comprehensive and accessible open data initiatives. Similarly, Latin America's mean HDI score of 0.74 is notably higher than Africa's 0.57, suggesting better overall human development outcomes in the Latin American region.

The standard deviation values suggest a wider spread of data in Latin America, for both NOD and HDI, compared to Africa. The kurtosis values are negative for both regions in terms of NOD, indicating light-tailed or less outlier-prone distributions. For HDI, Latin America shows positive kurtosis, indicating a heavy-tailed distribution, while Africa displays negative kurtosis. The skewness of NOD and HDI data in both regions indicates different distributions. For NOD, both regions are slightly positively skewed, with Latin America showing a slightly higher skewness. In contrast, HDI data shows a negative skewness in Latin America and a positive skewness in Africa.

Overall, the implication of these statistics for Hypothesis 2 is significant (see Table 5). The higher mean NOD and HDI values in Latin America suggest that the region's progress towards SDG3 targets is more pronounced than in Africa, potentially due to more effective or extensive open data initiatives. This affirms the link between open data and human development, reinforcing the need for enhanced open data strategies, especially in regions like Africa, to better progress towards global development goals.

The correlation coefficients between NOD initiatives and the HDI in Latin America and Africa reflect varying relationships (see Table 6). In Latin America, the correlation coefficient of 0.648 indicates a relatively strong positive relationship between NOD initiatives and the HDI. This suggests that an increase in the comprehensiveness and accessibility of open data initiatives is associated with higher levels of human development in this region. However, in Africa, the correlation coefficient of 0.279 suggests a weaker positive relationship between the NOD initiatives and the HDI. While there is a positive association, it is not as strong as in Latin America. An increase in open data initiatives might not lead to as substantial an increase in human development levels as in the Latin American context.

Table 5
Descriptive statistics for H2.

Indicator	Mean	Std. error	Median	Std. dev.	Sample variance	Kurtosis	Skewness	Min	Max
Latin America									
NOD	32.71	3.10	33.53	14.89	221.67	–0.96	0.25	7.97	58.04
HDI	0.74	0.02	0.75	0.08	0.01	0.75	–0.78	0.54	0.86
Africa									
NOD	19.97	1.39	20.16	6.35	40.39	–0.97	0.08	10.26	31.39
HDI	0.57	0.02	0.54	0.08	0.01	–0.47	0.63	0.45	11.91

Source: Authors' calculations.

Table 6
Pearson correlation for H2.

Variable	GDB	HDI
Latin America		
NOD	1	
HDI	0.64842	1
Africa		
NOD	1	
HDI	0.27900	1

Source: Authors' calculations.

These results highlight important regional variations in the relationship between open data initiatives and human development. It implies that while open data initiatives have a generally positive influence on human development, the strength of this relationship may differ based on regional contexts. It also emphasizes that other regional factors may also be influential in driving human development. These may include socio-political stability, levels of economic development, and access to education and healthcare. Therefore, it's crucial to take into account these contextual factors in formulating and implementing open data initiatives. In regions like Africa, where the correlation is weaker, it may be particularly important to focus on strengthening the enabling environment for open data to have a more substantial impact on human development.

The OLS regression results for Latin America and Africa suggest varying degrees of influence of NOD initiatives on the HDI (see Tables 7–8). For Latin America, the *R*-squared value is 0.42, meaning that around 42% of the variance in the HDI can be explained by the NOD. This is substantial and suggests that NOD initiatives have a significant impact on human development in the region. This interpretation is further supported by the *F*-statistics value of 15.23, and the *P*-value of 0.00, indicating that the model is statistically significant. The beta coefficient for NOD is 0.00, suggesting a slight but significant positive influence on HDI, as evidenced by a *t*-statistics value of 3.90 and an associated probability of 0.00. In contrast, for Africa, the *R*-squared value is only 0.08, indicating that NOD initiatives account for just 8% of the variance in the HDI. The *F*-statistics value is much lower at 1.60, and the probability of *F*-statistics is 1.60, suggesting the model is not statistically significant. The beta coefficient for NOD is also 0.00, indicating a negligible influence on HDI, a finding reinforced by a *t*-statistics value of 1.27 and a *P*-value of 0.22 indicating non-statistically significance of the model. By implication, NOD may not be a significant predictor of the dependent variable in the model.

Table 7
OLS regression for H2 — Latin America.

R-squared	Adjusted *R*-squared	*F*-statistics	Probability (*F*-statistics)	*P*-value
0.42	0.39	15.23	0.00	0.00

Variable	*β*	Std. error	*t*-statistics	Probability
NOD	0.00	0.00	3.90	0.00

Source: Authors' calculations.

Table 8
OLS regression for H2 — Africa.

R-squared	Adjusted *R*-squared	*F*-statistics	Probability (*F*-statistics)	*P*-value
0.08	0.03	1.60	1.60	0.22

Variable	*β*	Std. error	*t*-statistics	Probability
NOD	0.00	0.00	1.27	0.22

Source: Authors' calculations.

The comparison of these results indicates that the influence of NOD initiatives on HDI varies considerably between the two regions. In Latin America, open data initiatives seem to play a crucial role in human development, while in Africa, their impact is considerably less, pointing to the presence of other influential factors on human development.

These findings suggest that while promoting open data initiatives can contribute to human development, its impact may vary depending on regional contexts. Other factors, potentially including economic, social, or infrastructural variables, may also play significant roles, particularly in the context of Africa. Therefore, policymakers should adopt a comprehensive and context-specific approach, integrating open data initiatives with other development strategies to effectively enhance human development.

The Durbin–Watson test yielded a value of 2.071928278 (see Appendix B and C).

5. Discussion and conclusion

5.1. Discussion

The present research offers valuable insights into the role of NOD initiatives in shaping the HDI across different regions, substantiating existing literature that advocates the instrumental role of open data in fostering human development (Alamsyah et al., 2018; Husein et al., 2015; Kamil and Pratama, 2020; Virkar and Viale Pereira, 2018). Our study underscores the positive influence of open data initiatives on human development but also elucidates that the extent of this impact is variable across regions, with more pronounced effects observed in Latin America compared to Africa.

In Latin America, we identified a moderate positive correlation (0.648) between NOD and HDI, aligning with previous research demonstrating the potential of open data in enhancing transparency, accountability, and citizen engagement, ultimately leading to improved human development outcomes (Kamil and Pratama, 2020; Park and Gil-Garcia, 2017; Repkova Stofkova and Stofkova, 2020; Saxena and Muhammad, 2018; Virkar and Viale Pereira, 2018). The regression analysis further corroborates these findings, indicating that NOD initiatives can explain around 39% of the variation in HDI for this region, as suggested by the adjusted R-squared value of 0.39.

Contrastingly, in Africa, our study revealed a weaker positive correlation (0.279) between NOD and HDI, suggesting that the influence of open data initiatives on human development is comparatively less in this region. This conclusion is supported by the regression analysis, wherein the relationship between NOD and HDI was not statistically significant, with an adjusted R-squared value of a mere 0.03, implying that other factors predominantly drive human development in Africa. These findings resonate with previous ones (Algemili, 2016; Lineker and Runeson, 2020; Virkar and Viale Pereira, 2018), highlighting the unique challenges in implementing open data initiatives in Africa, such as limited infrastructural resources, scarcity of technical expertise, and institutional capacity issues.

While our findings endorse the significance of open data in advancing human development, they emphasize the necessity for region-specific strategies and interventions, considering the varied regional impact of these initiatives. Policymakers, thus, need to take into account the region-specific constraints and leverage open data initiatives accordingly to maximize their potential in driving human development.

5.2. Conclusion

In alignment with the core aim of this research to explore the impact of NOD initiatives on human capital development, particularly with respect to achieving SDG3 targets, the study reveals compelling evidence of a significant relationship between open data initiatives and the HDI across Latin America and Africa. The findings notably confirm that open data initiatives can act as a significant driver for human development, thereby contributing to the attainment of SDG3 targets. Therefore, based on the study's empirical results, Hypothesis H1 is accepted.

However, when examining the influence of open data initiatives on HDI across different regions, we observe considerable variance. While Latin America shows a notable positive correlation, the correlation is less pronounced in the case of Africa, suggesting region-specific challenges and dynamics. This significant variance in the impact of open data initiatives across Latin America and Africa leads us to accept Hypothesis H2, which postulates that these differences explain the varied progress towards achieving SDG3 targets across regions.

In essence, our study underscores the critical role of NOD initiatives in driving human development and achieving SDG3 targets, but also points out the importance of contextualizing these initiatives based on the unique regional dynamics. Policymakers are therefore urged to consider these findings in their strategic development plans, emphasizing the optimal utilization of open data initiatives in a manner that caters to the specific needs of each region, to enhance human development outcomes globally.

The research findings demonstrate that in Latin America, open data initiatives have a more pronounced and positive correlation with the HDI, which is a promising outcome. It indicates that current open data policies in Latin America are generally effective in advancing human development and are contributing positively toward achieving SDG3 targets. For policymakers in this region, this can be seen as a validation of the efforts put into open data and should encourage further investment and expansion of such initiatives. On the other hand, the correlation between open data initiatives and HDI in Africa is less conspicuous. This suggests that while open data holds promise as a lever for human development, the initiatives are not as effective in Africa due to specific regional challenges that might include infrastructure, data literacy, or governance issues. Policymakers in Africa need to recognize these limitations and consider strategies to make open data initiatives more impactful, potentially learning from the successes of Latin America.

Both regions are under the purview of the SDGs, and this study illuminates a path for how each can better align their open data initiatives with SDG3. For Latin America, the task might be more about optimizing and scaling current efforts, while for Africa, the focus may need to shift toward overcoming the unique challenges preventing open data from having a more significant impact.

6. Recommendations

Based on the findings of this study, the following recommendations are proposed to maximize the impact of open data initiatives on human development and to achieve the SDG3 targets:

- *Strengthen open data initiatives.* Governments and stakeholders should invest in strengthening open data initiatives by improving data quality, ensuring data accessibility, and promoting data interoperability. This will enable individuals, organizations, and communities to utilize open data effectively for decision-making, innovation, and human development.
- *Build capacity and infrastructure.* Policymakers should prioritize capacity building and infrastructure development to support open data initiatives, particularly in regions where such initiatives have a less pronounced impact on human development. This may include investing in digital infrastructure, enhancing data literacy, and providing training and resources to local institutions and organizations.
- *Foster collaboration and partnerships.* Governments and stakeholders should collaborate and form partnerships with relevant organizations, including international organizations, non-governmental organizations, and the private sector, to promote the effective implementation of open data initiatives. This collaboration can help leverage resources, knowledge, and expertise to address challenges and capitalize on opportunities in different regions.
- *Adopt a context-specific approach.* Policymakers should design and implement open data initiatives based on the unique socio-economic, cultural, and institutional contexts of their respective regions. This will ensure that open data initiatives are tailored to address local challenges and opportunities, thereby enhancing their impact on human development.
- *Monitor and evaluate open data initiatives.* Governments and stakeholders should establish robust monitoring and evaluation systems to assess the impact of open data initiatives on human development over time. This will enable the identification of best practices, the adjustment of policies and strategies, and the continuous improvement of open data initiatives.
- *Conduct further research.* Researchers should continue to explore the relationship between open data initiatives and human development, with a focus on understanding the factors that mediate this relationship and the mechanisms through which open data contributes to human development. This research will provide valuable insights for policymakers and stakeholders to design and implement effective open data initiatives that promote human development and achieve the SDG3 targets.

7. Limitation of the study

This research centered on the relationship between open data initiatives and human capital development in the context of SDG3 targets, acknowledges several limitations. The study uses secondary data for regression analysis, which could lead to issues in data accuracy, consistency, and completeness. The adopted cross-sectional approach provides only a snapshot of the relationship at a particular point in time, potentially missing out on long-term impacts of open data initiatives. While correlations are established through regression analysis, causality cannot be definitively determined, as the relationship could be influenced by unaccounted confounding factors. Moreover, the study's generalizability may be restricted due to varying political, economic, social, and cultural contexts across different regions. Also, given the broad SDG3 targets and the complex relationship between open data and human development, some aspects may remain unexplored. Despite these limitations, the study contributes valuable insights to the existing literature, aiding in optimizing open data initiatives for improved health outcomes and human development.

References

Alamsyah, A., Gustyana, T. T., Fajaryanto, A. D., & Septiafani, D. (2018). *Open data analytical model for Human Development Index optimization to support government policy.* arXiv:1809.00189. https://doi.org/10.48550/arXiv.1809.00189

Algemili, U. A. (2016). Outstanding challenges in recent open government data initiatives. *International Journal of E-Education, e-Business, e-Management and e-Learning*, *6*(2), 91–102. https://doi.org/10.17706/ijeeee.2016.6.2.91-102

Chen, Y. (2016). Spatial autocorrelation approaches to testing residuals from least squares regression. *PLoS ONE*, *11*(1), e0146865. https://doi.org/10.1371/journal.pone.0146865

Constantine, N. A. (2012). Regression analysis and causal inference: Cause for concern? *Perspectives on Sexual and Reproductive Health*, *44*(2), 134–137. https://doi.org/10.1363/4413412

Criado, J. I., Jimenez, C., & Alcaide-Munoz, L. (2021). Open data portals development in city councils. An empirical analysis based on structural factors in Spain. In *DG.O2021: The 22nd Annual International Conference on Digital Government Research* (pp. 132–141). New York: Association for Computing Machinery. https://doi.org/10.1145/3463677.3463737

Analyttica Datalab (2021). Understanding Durbin–Watson test. *Medium*, August 4. https://medium.com/@analyttica/durbin-watson-test-fde429f79203.

de Beer, J. (2017). Open innovation in development: Integrating theory and practice across open science, open education, and open data. *Open AIR Working Paper*, No 3/17. https://doi.org/10.2139/ssrn.3008675

Ding, C. S. (2006). Using regression mixture analysis in educational research. *Practical Assessment, Research, and Evaluation*, *11*(11), 11. https://doi.org/10.7275/wgt2-b390

Ekundayo, T. (2021). Leveraging national data governance to drive economic change. *Organization Leadership and Development Quarterly*, *4*(1), 30–46.

Ekundayo, T., Bhaumik, A., & Chinoperekweyi, J. (2023a). Identifying the core data governance framework principle: A framework comparative analysis. *Organization Leadership and Development Quarterly, 5*(1), 30–53.

Ekundayo, T., Bhaumik, A., Chinoperekweyi, J., & Khan, Z. (2023b). The impact of open data implementation on entrepreneurship ability in Sub-Saharan Africa. *Human Behavior and Emerging Technologies*, *2023*, 1–11. https://doi.org/10.1155/2023/7583550

Ekundayo, T., & Isaac, O. (2023). Open data: A national data governance strategy for open science and economic development—A case study of the United Arab Emirates. *Emirati Journal of Business, Economics and Social Studies*, *1*(2), 98–109. https://doi.org/10.54878/ejbess.208

Fan, X., Seshadri, A., & Taber, C. R. (2022). Estimation of a life-cycle model with human capital, labor supply and retirement. *NBER Working Paper*, No. w29905. https://doi.org/10.3386/w29905

Hess, A. S., & Hess, J. R. (2017). Linear regression and correlation. *Transfusion*, *57*(1), 9–11. https://doi.org/10.1111/trf.13928

Husein, I. G., Danar Sunindyo, W., Bahawares, R., Nainggolan, Y., & Akbar, S. (2015). Open data strategy for enhancing the productivity and competitiveness of fishery SMEs in Indonesia. In *2015 International Conference on Electrical Engineering and Informatics (ICEEI)* (pp. 490–495). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ICEEI.2015.7352550

Huston, P., Edge, V. L., & Bernier, E. (2019). *Reaping the benefits of open data in public health*. *Canada Communicable Disease Report, 45*(10), 252–256. https://doi.org/10.14745/ccdr.v45i10a01

Kamil, M., & Pratama, M. I. (2020). Openness, digital and human development: Case study of ASEAN countries. In *Proceedings of the 2nd International Research Conference on Economics and Business IRCEB* (vol. 1, pp. 187–192). Universitas Negeri Malang, Indonesia. https://doi.org/10.5220/0008785701870192

Latifah, L., Setiawan, D., Aryani, Y. A., Sadalia, I., & Al Arif, M. N. R. (2022). Human capital and open innovation: Do social media networking and knowledge sharing matter? *Journal of Open Innovation: Technology, Market, and Complexity*, *8*(3), 116. https://doi.org/10.3390/joitmc8030116

Lineker, J., & Runeson, P. (2020). Collaboration in open government data ecosystems: Open cross-sector sharing and co-development of data and software. In G. Viale Pereira et al. (Eds.), *Electronic government. EGOV 2020. Lecture notes in computer science* (vol. 12219, pp. 290–303). Cham: Springer. https://doi.org/10.1007/978-3-030-57599-1_22

Magalhres, J., Hartz, Z., Antunes, A., & Martins, M. do R. O. (2017). An overview of the open science in times of big data and innovation to global health. *International Journal of Innovation*, *5*(3), 270–288. https://doi.org/10.5585/iji.v5i3.219

Martin, E. G., & Begany, G. M. (2017). Opening government health data to the public: Benefits, challenges, and lessons learned from early innovators. *Journal of the American Medical Informatics Association*, *24*(2), 345–351. https://doi.org/10.1093/jamia/ocw076

Meurs, M., Seidelmann, L., & Koutsoumpa, M. (2019). How healthy is a "healthy economy"? Incompatibility between current pathways towards SDG3 and SDG8. *Globalization and Health*, *15*(1), 1–13. https://doi.org/10.1186/s12992-019-0532-4

Njoku, J. U., & Onyegbula, J. C. (2017). Human capital development as a strategy for sustainable development in the Nigerian education system. *African Research Review*, *11*(2), 178–189. https://doi.org/10.4314/afrrev.v11i2.13

Park, S., & Gil-Garcia, J. R. (2017). Understanding transparency and accountability in open government ecosystems: The case of health data visualizations in a state government. In *Proceedings of the 18th Annual International Conference on Digital Government Research (dg.o' 17)* (pp. 39–47). New York: Association for Computing Machinery. https://doi.org/10.1145/3085228.3085318

Repkova Stofkova, K., & Stofkova, J. (2020). Use of open data in the development of the digital economy in the knowledge society in the era of globalization. *SHS Web of Conferences*, *74*, 03008. https://doi.org/10.1051/shsconf/20207403008

Saunders, M. N. K., Lewis, P., & Thornhill, A. (2007). Understanding research philosophy and approaches to theory development. In *Research methods for business students* (8th ed., pp. 128–171). ParkHarlow: Pearson Education.

Saxena, S., & Muhammad, I. (2018). The impact of open government data on accountability and transparency. *Journal of Economic and Administrative Sciences*, *34*(3), pp. 204–216. https://doi.org/10.1108/jeas-05-2017-0044

Sebubi, O., Zlotnikova, I., & Hlomani, H. (2020). Open data for sustainable development on a knowledge-based economy: The case of Botswana. *Data Science Journal*, *19*(1), 44. https://doi.org/10.5334/DSJ-2020-044

Sedgwick, P. (2013). Correlation versus linear regression. *BMJ, 346*, f2686. https://doi.org/10.1136/bmj.f2686

Shabbir, T., & Nadeemullah, M. (2020). Impact of "open data" and its effectiveness for Pakistan social issues: Learning from the UK experience. *Pakistan Perspectives*, *25*(1), 253–272.

Školudová, J. (2016). Human capital management: Monitoring of the key employees in organizations in the Czech Republic. In *28th International Business Information Management Association Conference* (pp. 2537–2541). Norristown: IBIMA,

Turner, P. (2020). Critical values for the Durbin–Watson test in large samples. *Applied Economics Letters*, *27*(18), 1495–1499. https://doi.org/10.1080/13504851.2019.1691711

Virkar, S., & Viale Pereira, G. (2018). Exploring open data state-of-the-art: A review of the social, economic and political impacts. In P. Parycek et al. (Eds.), *Electronic government. EGOV 2018. Lecture notes in computer science* (vol. 11020, pp. 196–207). Cham: Springer. https://doi.org/10.1007/978-3-319-98690-6_17

Appendix A. Global dataset

Country	NOD (x)	HDI (y)
Albania	38.3108301	0.796
Angola	10.5644908	0.586
Argentina	50.4105292	0.842
Armenia	44.5520431	0.759
Australia	55.4746035	0.951
Azerbaijan	21.8086519	0.745
Bahamas	23.9256527	0.812
Bahrain	21.9810401	0.875
Bangladesh	23.7534615	0.661
Belarus	19.4686636	0.808
Belize	24.3902938	0.683
Benin	14.3655633	0.525
Bolivia (Plurinational State of)	21.9714181	0.692
Botswana	20.1584788	0.693
Brazil	58.0474901	0.754
Bulgaria	49.6532605	0.795
Burkina Faso	22.5402645	0.449
Cambodia	13.1575595	0.593
Cameroon	24.1453459	0.576
Canada	60.8230052	0.936
Chile	52.8939430	0.855
China	39.8227573	0.768
Colombia	53.7692965	0.752
Costa Rica	34.4980051	0.809
Côte d'Ivoire	19.8295847	0.550
Croatia	47.9013403	0.858
Czechia	45.0215017	0.889
Denmark	58.1908817	0.948
Dominican Republic	35.1617520	0.767
Ecuador	34.5730924	0.740
Egypt	21.8340247	0.731
El Salvador	13.4462315	0.675
Estonia	67.3547208	0.890
Finland	54.5010533	0.940
France	66.2249337	0.903
Gambia	20.4943385	0.500
Georgia	40.2500405	0.802
Germany	58.0693664	0.942
Ghana	27.6630628	0.632
Greece	36.6018968	0.887
Guatemala	18.7529931	0.627
Guyana	11.2251695	0.714
Haiti	7.9717546	0.535
Honduras	24.9406319	0.621
India	46.6843411	0.633
Indonesia	40.2350731	0.705
Ireland	46.0489305	0.945
Israel	42.0811688	0.919
Italy	56.5443001	0.895
Jamaica	30.9746313	0.709
Jordan	22.1906687	0.720
Kazakhstan	41.6609239	0.811
Kenya	25.7182937	0.575

(continued on next page)

Appendix A *(continued)*

Country	NOD (x)	HDI (y)
Kyrgyz Republic	23.5012886	0.692
Latvia	49.1745634	0.863
Liberia	17.1995856	0.481
Lithuania	37.2995525	0.875
Malawi	14.5983259	0.512
Malaysia	41.5507405	0.803
Malta	36.5436621	0.918
Mexico	50.6439854	0.758
Mongolia	32.8294645	0.739
Morocco	12.3720943	0.683
Mozambique	10.2613326	0.446
Namibia	18.8841260	0.615
Nepal	18.9394142	0.602
Netherlands	54.0301386	0.941
New Zealand	65.5636099	0.937
Nigeria	24.2595673	0.535
Oman	14.1401969	0.816
Panama	34.5874656	0.805
Paraguay	33.5248256	0.717
Peru	37.6498035	0.762
Philippines	34.0434394	0.699
Portugal	41.9288291	0.866
Qatar	22.2145184	0.855
Romania	43.0174259	0.821
Russian Federation	41.6516827	0.822
Rwanda	24.8139586	0.534
Saint Lucia	21.3100328	0.715
Saudi Arabia	29.0411437	0.875
Senegal	12.0579508	0.511
Slovakia	50.8822867	0.848
South Africa	30.3563192	0.713
Spain	55.8205998	0.905
Sri Lanka	16.347382	0.782
Sweden	42.7811905	0.947
Tajikistan	12.2370099	0.685
Thailand	41.7482278	0.800
Togo	14.5612244	0.539
Trinidad and Tobago	22.4177916	0.810
Tunisia	23.0695951	0.731
Turkmenistan	6.4075570	0.745
Uganda	31.3989151	0.525
Ukraine	55.4857581	0.773
United Arab Emirates	26.6912191	0.911
United Kingdom of Great Britain and Northern Ireland	64.5356916	0.929
United States of America	68.0199151	0.921
Uruguay	55.2306432	0.809
Uzbekistan	31.7442048	0.727
Vietnam	33.3245030	0.703

Note: Durbin–Watson test of auto-correlation. Formula for Durbin–Watson: $d = \sum_{t=2}^{T}(u_t - u_{t-1})^2 / \sum_{t=1}^{T} u_t^2$, where: T — the total number of observations; u_t — the t-th residual from the regression model.
Calculations: Sum of squared difference of the residual / Sum of squared residuals, where: Sum of squared difference of the residual = 1.876988276; Sum of squared residuals = 0.905913731. Durbin–Watson = 2.071928278.
Source: Authors' calculations.

Appendix B. Latin America dataset

Country	NOD (x)	HDI (y)
Argentina	50.41052920	0.842
Bahamas	23.92565272	0.812
Belize	24.39029375	0.683
Bolivia (Plurinational State of)	21.97141810	0.692
Brazil	58.04749012	0.754
Chile	52.89394295	0.855
Colombia	53.76929648	0.752
Costa Rica	34.49800510	0.809
Dominican Republic	35.16175198	0.767
Ecuador	34.57309243	0.740
El Salvador	13.44623151	0.675
Guatemala	18.75299306	0.627
Guyana	11.22516953	0.714
Haiti	7.97175461	0.535
Honduras	24.94063186	0.621
Jamaica	30.97463130	0.709
Mexico	50.64398542	0.758
Panama	34.58746564	0.805
Paraguay	33.52482560	0.717
Peru	37.64980354	0.762
Saint Lucia	21.31003275	0.715
Trinidad and Tobago	22.41779158	0.810
Uruguay	55.23064323	0.809

Note: Durbin–Watson test of auto-correlation. Formula for Durbin–Watson: $d = \sum_{t=2}^{T}(u_t - u_{t-1})^2 / \sum_{t=1}^{T} u_t^2$, where: T — the total number of observations; u_t — the t-th residual from the regression model.
Calculations: Sum of squared difference of the residual / Sum of squared residuals, where: Sum of squared difference of the residual = 0.156888906; Sum of squared residuals = 0.075758319. Durbin–Watson = 2.070913249.
Source: Authors' calculations.

Appendix C. Africa dataset

Country	GDB (x)	HDI (y)
Angola	10.56449075	0.586
Benin	14.36556330	0.525
Botswana	20.15847878	0.693
Burkina Faso	22.54026453	0.449
Cameroon	24.14534591	0.576
Côte d'Ivoire	19.82958469	0.550
Gambia	20.49433847	0.500
Ghana	27.66306284	0.632
Kenya	25.71829365	0.575
Liberia	17.19958557	0.481
Malawi	14.59832587	0.512
Morocco	12.37209429	0.683
Mozambique	10.26133256	0.446
Namibia	18.88412595	0.615
Nigeria	24.25956727	0.535
Rwanda	24.81395861	0.534
Senegal	12.05795081	0.511
South Africa	30.35631918	0.713
Togo	14.56122436	0.539
Tunisia	23.06959507	0.731
Uganda	31.39891509	0.525

Note: Durbin–Watson test of auto-correlation. Formula for Durbin–Watson: $d = \sum_{t=2}^{T}(u_t - u_{t-1})^2 / \sum_{t=1}^{T} u_t^2$, where: T — the total number of observations; u_t — the t-th residual from the regression model.
Calculations: Sum of squared difference of the residual / Sum of squared residuals, where: Sum of squared difference of the residual = 0.355572168; Sum of squared residuals = 0.127168312. Durbin Watson = 2.796075231.
Source: Authors' calculations.