Finance Research and the UN Sustainable Development Goals – an analysis and forward look

Yang Su[a]*, Brian M. Lucey [a,b,c,d], Ashish Kumar Jha[a]*

[a] *Trinity Business School, Trinity College Dublin, Dublin 2, Ireland*

[b] *Abu Shabi University College of Business, Abu Dhabi*

[c] *Distinguished Research Fellow, Institute of Business Research, University of Economics Ho Chi Minh City, 59C Nguyen Dinh Chieu, Ward 6, District 3, Ho Chi Minh City, Vietnam*

[d] *Institute for Industrial Economics, Jiangxi University of Economics and Finance, 169, East Shuanggang Road, Xialuo, Changbei District 330013 Nanchang, Jiangxi, China*

Corresponding Author: **Yang Su** *suya@tcd.ie*

Abstract

This study conducts a comprehensive analysis of the interplay between the United Nations Sustainable Development Goals (UN SDGs) and scholarly output in financial journals from 2010 to 2022. Utilizing a framework that combines fuzzy matching, bibliometric analysis, and topic modelling techniques, the research traces the evolution of finance research and its alignment with the UN SDGs. The findings demonstrate a focus within finance research on Economic Growth (Goal 8) and Peace and Justice (Goal 16), while also identifying areas that warrant further scholarly attention. Key discoveries include: (1) the enhanced effectiveness of fuzzy matching when augmented by RAKE (Rapid Automatic Keyword Extraction) for feature extraction, particularly in the context of complex datasets; (2) the mismatch between actual sustainability progress and SDG related research influence in the USA China and India; and (3) significant differences across levels of the ABS ranking in terms of alignment with UN SDG's with ABS 4 journals demonstrating a comparatively low output and influence.

Keywords: UN Sustainable Development Goals; Fuzzy Matching Analysis; Topic Modeling; Financial Academic Trends

JEL Code: Q01, O10, C38, A11

---

1. Introduction

In 2015, the United Nations unveiled the "2030 Agenda for Sustainable Development," also known as the United Nations Sustainable Development Goals (UN SDGs), which includes 17 objectives. The fundamental purpose of these objectives is to enhance health and educational outcomes, mitigate disparities, foster economic expansion, and address the implications of climate change, while also prioritizing the conservation of marine ecosystems and forests [United Nations, 2015]. The UN SDGs evolved from the Millennium Development Goals (MDGs). Compared to the MDGs, the UN SDGs place a heightened emphasis on the significance of sustainable development [Sachs, 2012]. Given the importance of sustainable development, the United Nations Sustainable Development Goals (UN SDGs) have increasingly garnered attention and recognition from the global community, especially in academia [Oestreich, 2023]. In a world faced with complex environmental, social, and economic challenges, the establishment of the UN SDGs marked a shift in our global trajectory, representing a universally agreed call to action across various sectors, including finance [United Nations, 2015]. They urge various sectors and disciplines to align their practices and research with these targets. Facing this, some argue that the financial sector has undergone a paradigm shift, with an increasing emphasis on integrating environmental, social, and governance (ESG) factors into its fabric. This change can be partly attributed to evidence of the financial benefits associated with sustainable investing [Renneboog et al., 2008]. However, Atz et al. [2023] presents a nuanced view, suggesting that sustainability may not always lead to enhanced financial performance, thereby adding complexity to the discourse on ESG investing. Furthermore, the demand for responsible operations and transparent sustainability disclosures from stakeholders, ranging from investors to consumers, has intensified [Cheng et al., 2023]. [Edmans, 2023] offers a comprehensive overview of the current state and challenges of ESG in the financial landscape, illustrating the evolving nature of sustainability considerations in finance. Additionally, de Souza Barbosa et al. [2023] highlights the positive impacts of ESG criteria integration on corporate non-financial performance, reinforcing the multifaceted effects of ESG practices in aligning with the UN SDGs.

As industry practices have adapted to accommodate evolving sustainability imperatives, it becomes important to explore the extent to which academic finance research has aligned with the substantial sustainability agenda introduced by the SDGs. Scholarly publications

2

are instrumental in moulding dialogue, steering pedagogical methodologies, and guiding the direction of industry practices. In a similar vein, Bebbington and Unerman [2018] have underscored the role of academic investigations into sustainable development, emphasizing their potential to shape policy decisions, notably in the accounting sector. Analyzing this congruence, or the potential divergence, provides insights into the overarching direction of the finance discipline, elucidating its predominant areas of emphasis, existing gaps, and prospective paths, as supported by Beck et al. [2006].

In mapping out our methodological approach, we employed a multi faceted framework. Articles from financial journals underwent an in-depth review. For textual analysis, we amalgamated keyword extraction with Fuzzy Matching techniques to pinpoint specific themes and narratives pertinent to the SDGs. The integration of Fuzzy Matching, known for its precision in text classification, bolsters the fidelity of our thematic identification [Chaudhuri et al., 2003]. It is important to emphasize that instead of establishing a new index, our study offers an *optimized keyword list* tailored for Fuzzy Matching. This list provides a potentially useful tool for scholars delving into SDG-related financial research. In academia, the emphasis on aligning with the Sustainable Development Goals (SDGs) is becoming increasingly important Alcántara-Rubio et al. [2022]. Within this context, the Fuzzy Match algorithm stands out as a ueful tool. Its proficiency in categorizing large datasets, such as the substantial corpus of financial articles pertinent to this study, is noteworthy. Fuzzy Match significantly aids in pinpointing articles that are relevant to the UN SDGs, enabling a more targeted and thorough investigation into how the academic finance community is addressing these vital global aims. Also drawing on commonly used literature research methods, the study incorporates Bibliometrics to methodically analyze and map out the academic landscape (Ball, 2017; Mbarki et al., 2023; Lucey et al., 2023; Pandey et al., 2023).

Bibliometric methods have become common in finance research. A comprehensive guide to bibliometric and scientometric research approaches is contained in Donthu et al. [2021], while Khan et al. [2022] offer a detailed a bibliometric analysis of finance bibliometric, analysing over 120 papers as at that time.

This, coupled with our findings, enables us to chart clear trajectories at the intersection of finance and sustainable development. By synthesizing these insights, the research aims to offer forward-looking recommendations, highlighting areas within finance and sustainability that warrant increased scholarly focus. The paper is organized as follows: Section One

3

provides an introduction, section Two details the data utilized. Thereafter, section Three outlines our methodology, section Four the empirical analysis and section Five concludes and offers insights for future research.

2. Data

2.1. Data Collection and Data Cleaning

We confine here our study to finance journals indexed under the Academic Journal Guide 2021(ABS 21) list. This list is in widespread use as a guide to management journal qualityliterature(pbryce2020journal ; pwalker2019influences). We used the comprehensive Scopus database to obtain scholarly articles from the 113 finance journals categorized under ABS 2021. This, spanning the years 2010 to 2022, yielded an initial set of 60,504 articles.

We then undertook a data cleaning process. This first involved excluding articles with irregular or missing abstracts. Post-cleaning, our dataset was streamlined to 53,792 articles, marking a retention rate of 88.90%. This curated collection, embracing both articles and their abstracts from the finance domain, forms the bedrock of our analyses.

2.2. Data Processing

In the data processing phase of this study, we initially experimented with eight different methodologies to determine the relevance of these 53,792 financial articles to the UN SDGs. This involved a trial process, with both computational assessment and manual validation by experts in sustainable finance applied to a random sample of 400 articles. The objective was to ascertain the most effective model for discerning the articles' alignment with the UN SDGs.

[PLEASE INSERT Figure 1 HERE]

Following this testing phase, the study successfully identified an optimal model that combined the Fuzzy Match technique with specific keywords. Upon applying this approach to the entire dataset of 53,792 articles, 10,031 were identified as relevant to the UN SDGs, representing 18.64% of the total, as shown in Figure 1. Among them, 597 articles were classified to be related to 2 or 3 different UNSDGs Goals. This approach underscores both the effectiveness of combining computational methods with expert insights to efficiently filter and categorize a large corpus of academic literature in line with sustainability goals and the possible lack of overall alignment.

4

3. Methodology

We use a three step approach. Initially, a Fuzzy Match approach was employed for data filtering, involving a comparative analysis of eight different methodologies. Following a manual review, the most effective method was determined, specifically the use of keywords identified through the Rake algorithm as the source for Fuzzy Match. The research then progressed to a detailed analysis phase, utilizing Bibliometric and Topic Modelling techniques to conduct an in-depth content analysis and empirical examination of the literature.

3.1. Fuzzy Match

In this research, the Fuzzy Match approach is pivotal for categorizing abstracts from financial journals, assessing their relevance to the United Nations Sustainable Development Goals (UN SDGs). This method, adept at identifying non-exact matches in data, is particularly suited for handling textual data's inherent variations and inconsistencies. For further details on the fuzzy matching methodology and an introduction to the model used in this article, please refer to Appendix C.

When classifying text, keyword classification is a very common method. However, Li et al. [2017] indicate that an approach based on fuzzy logic is more effective for text classification compared to the keyword method, offering higher accuracy than other methods. The adaptability of Fuzzy Match algorithms to different dataset requirements makes it an useful tool for categorization challenges. The technique's ability to enhance data quality and ensure the precision and reliability of the categorization process is a significant factor in its selection for this research.

Fuzzy match usually inputs the original text as the training source. Therefore, this study first selected the short original text of the UN SDGs (including 17 items broader goals text and subcategories goals text) as the input of Model 1, and the long original text(only containing UN SDGs broader goals text) as the input of Model 2. However, original text of UN SDGs might contain an excess of descriptive words, potentially impacting the effectiveness of the match. To address this, three other commonly used text feature extraction methods were selected to simplify the long and short original texts of UN SDGs respectively. By extracting five to fifteen key feature words to do the pre-processing for model training. There are some studies that found that Fuzzy Keyword Match is effective to some extent, so this study will conduct experiments (Ahsan et al., 2017; Song et al., 2009;

5

Wang et al., 2014). This extraction employed traditional TF-IDF (Term Frequency-Inverse Document Frequency) and RAKE (Rapid Automatic Keyword Extraction) algorithms. TF-IDF algorithm weighs the frequency of a word in a document against its frequency across all documents, thereby highlighting words that are unique to a specific document. Rake, on the other hand, identifies key phrases in a text based on the frequency of word appearance and its co-occurrence with other words. Additionally, the innovative GPT-4.0 model was incorporated. By instructing GPT-4.0 to identify "5-15 words that best represent the essence of the text", a novel approach to feature extraction was explored. Subsequent analysis of these methods' accuracy in the Fuzzy Match process allowed for the identification of the most effective strategy. The specific content and accuracy of these eight models are placed in Table 1 and Table 2.

[PLEASE INSERT Table 1 and Table 2 HERE]

The study involved a thorough comparison of eight distinct Fuzzy methodologies to ascertain the most effective Fuzzy Match approach. This study builds upon existing concepts, specifically applying, and refining the Fuzzy Match approach previously proposed. It introduces a similarity function that aims to overcome the constraints of conventional similarity measures, thereby contributing to the advancement of a more efficient algorithm for fuzzy matching, as initially outlined in the works of Cayrol et al. [1982] and Chaudhuri et al. [2003].

A flowchart in Appendix C demonstrates the methodology, from the training of the classifier with SDG-specific documents to the processing of abstracts and determination of their relevance based on similarity scores.

By adopting these advanced fuzzy match techniques, the study offers a nuanced understanding of the trajectory of finance research within the context of the UN SDGs. This methodological choice highlights the importance of sophisticated data categorization techniques in academic research, especially in large-scale data analyses aimed at extracting deep insights from complex datasets.

This study found that applying Fuzzy Match as a classification method and using Rapid Automatic Keyword Extraction (Rake) keyword list to long text will improve the accuracy of classification, as the model 8 shown in Table 1 and Table 2.

3.2. Bibliometric

Bibliometric analysis, a systematic and quantitative approach, was utilized in this study to evaluate literature. This method assesses various bibliometric indicators such as publication trends, citation analysis, authorship patterns, and co-authorship networks, offering a data-driven, objective way to map the intellectual structure of finance research [Donthu et al., 2021].

Widely applied in finance, bibliometric analysis is adept at revealing emerging research areas, aligning perfectly with the aim of this study to integrate financial research with sustainable development goals. One paper indicates that its effectiveness in highlighting novel academic themes underscores its suitability for this research [Mbarki et al., 2023].

In this study, employing computational methods derived from R tools [Aria and Cuccurullo, 2017], we recalibrated our approach for the bibliometric analysis due to the extensive volume of data. Our focused analysis reevaluates publication trends, geopolitical dynamics, and identifies the top five publishing countries, integrating this with ABS rankings. The study, while comprehensive, does not extend into exploring author relationships, as our primary objective is to establish a high-level mapping between SDG-focused finance academia and broader publication patterns.

3.3. Topic Modelling

Latent Dirichlet Allocation (LDA) was created by Blei et al. [2003] to uncover hidden thematic structures in large volumes of text data. In this study, LDA was employed to extract and analyze topics from financial literature spanning from 2010 to 2022. LDA's proficiency in discovering latent topics in extensive datasets, typical of financial journals, made it a fitting choice. This probabilistic model, which views documents as mixtures of topics and topics as mixtures of words, differs from bibliometric methods that focus on word frequency. It is adept at revealing hidden thematic structures, essential for elucidating the relationship between finance research and UN SDGs.

The implementation of LDA involved determining the optimal number of topics by assessing the perplexity of various models, resulting in the identification of seven key domains. This facilitated trend analysis and the creation of a topic similarity matrix, providing insights into the interconnections and parallels between the topics. This approach not only enhanced the comprehension of evolving themes in finance research but also highlighted its alignment with global sustainability strategies and shifting academic priorities [Li et al.,

7

2017]. The choice of LDA, a method widely applied in finance for analyzing complex data, aligns the study with established practices in the field, ensuring a thorough and relevant exploration of the subjects in the context of the UN SDGs. For further details on the LDA methodology and an introduction to the model used in this article, please refer to Appendix D.

4. Empirical Results

4.1. Overall Analysis

4.1.1. Analysis of Finance-Related Research in the Context of UN SDGs

[PLEASE INSERT Figure 2 and Figure 3 HERE]

The analysis presented in Figure 3 and Figure 2 represents an "Evolution of Finance Research" concerning the UN SDGs. These figures, displaying the percentage distribution of research focus, illustrate a dynamic shift in the finance community's priorities from 2010 to 2022. There is a noticeable increase in attention towards Goal 7 (Affordable and Clean Energy) and Goal 13 (Climate Action), alongside a gradual growth in research related to Goal 5 (Gender Equality). Despite a proportional decrease in focus, Goal 8 (Decent Work and Economic Growth) maintains a significant presence in terms of overall research interest. This study's findings align with Kumar et al. [2022] research on sustainable finance research trends , which emphasizes key topics in sustainable finance are about socially responsible investing, climate financing, and energy financing. These key topics match with our statistical results, suggesting a growing consensus in the finance community about the critical nature of these sustainable finance topics.

[PLEASE INSERT Figure 4 and Figure 5 HERE]

[PLEASE INSERT Table 3 HERE]

Figure 4, and Table 3 present the actual counts of publications and citations, offering a different perspective. From a quantitative standpoint, Goal 8(Economic Growth) emerges as

8

the most extensively discussed and cited, with its primary objectives aligning closely with the majority of finance-related literature. It significantly outpaces other goals in terms of research output. Following Goal 8 (Economic Growth), Goal 16 (Peace, Justice, and Strong Institutions) occupies a prominent place in the finance literature, both in citations and publications. A notable finding is the extensive citation of papers related to Goal 5 (Gender Equality) and Goal 16 (Peace and Justice). Despite publication counts on gender topics being roughly equivalent to those on climate and urban issues, the citation numbers are two to three times higher. This indicates a deep discussion on gender topics, with each research piece carrying substantial impact, potentially extending its influence beyond the finance sector into other areas. Goal 17 (Partnerships for the Goals) is represented by a significant volume of research, contrary to a notion of gradual increase in focus. This volume underlines the finance sector's recognition of the importance of collaborative efforts, including public-private partnerships and international financial initiatives, in achieving sustainable development.

The limited exploration of certain goals, such as Goals 14 (Life Below Water), 15 (Life on Land), and 6 (Clean Water and Sanitation), clearly reflects their relatively lower relevance within the traditional scope of financial research. This may change with the advent of papers such as Karolyi and Tobin-de la Puente [2023] and Flammer et al. [2023]. However, this also underscores the need for increased attention to these areas, suggesting an opportunity for integrating environmental finance with traditional financial research to create innovative financial models and strategies for sustainable development.

Figure 5 provides a differential analysis of citation versus publication percentages for UN SDGs in finance research. The calculation behind each bar is the citation percentage minus the publication percentage for each specific SDG. Bars above the zero line represent a positive differential, where the citations for an SDG outweigh the publications, suggesting that these areas receive higher academic impact or attention relative to the number of papers published.

Positive differentials, shown as blue bar, are observed for SDGs like Goal 16 (Peace, Justice, and Strong Institutions), as indicated by bars extending above the zero line, signifying a higher citation rate compared to publication rate. This suggests these topics not only have a solid presence in research publications but also possess substantial academic influence, resonating through subsequent scholarly works. It may even be possible to impress other fields besides finance.

9

On the other hand, negative differentials, seen in SDGs such as Goal 10 (Reduced Inequalities) and Goal 17 (Partnerships), are represented by bars dipping below the zero line. These reflect a lower citation rate relative to the publication rate, indicating that despite their research popularity, they garner fewer citations than might be expected. This suggests an undercitation in these areas, pointing to potential underrecognized yet important research within the finance literature that could benefit from increased scholarly engagement and citation.

In summary, finance-related research demonstrates a pronounced alignment with certain UN Sustainable Development Goals, notably Goal 8 (Decent Work and Economic Growth), which consistently garners significant interest among scholars. Concurrently, there is a attention towards Goal 7 (Affordable and Clean Energy) and Goal 13 (Climate Action), indicating an expanding research focus that extends beyond traditional economic parameters. Moreover, the influence of research on Goal 5 (Gender Equality) is increasingly notable, not only within the realm of finance but potentially impacting broader domains beyond finance. This trend underscores a shift towards perhaps a more holistic approach in finance research, recognizing the critical interplay between economic growth, environmental sustainability, and social equity. Such a shift calls for a more inclusive research agenda, one that embraces both immediate economic concerns and broader sustainability challenges, thereby contributing to a more comprehensive understanding of sustainable development and fostering a more balanced and sustainable global economic system.

4.1.2. Geographical Trend Analysis

[PLEASE INSERT Table 4 HERE]

As shown in Table 4, from 2010 to 2022, the statistics of the top ten countries in terms of number of publications can be expressed in this column. This column can express the academic performance ranking of the country. At the same time, the table shows the rankings of these ten countries in two different sustainable development-related indices. Reflecting on the nexus between financial research and sustainability, a juxtaposition of our bibliometric data against the Environmental Performance Index (EPI)[1] , compiled by Yale University (Wolf et al., 2022), and the Sustainable Development Goals Index (SDGI)[2], prepared by the

---

[1]https://epi.yale.edu/epi-results/2022/component/epi
[2]https://dashboards.sdgindex.org/rankings

10

SDSN and the IEEP((IEEP), 2023), reveals nuanced correlations. Assessments of sustainable development progress, as reflected in the SDGI and the EPI, indicate that Germany, France, the United Kingdom, and Spain are essentially leading the way, with their rankings closely aligning with their academic performance rankings. This interplay suggests that in these countries, financial research may be contributing to, or at least progressing in tandem with, advances in sustainable development. These rankings not only mirror policy initiatives but also demonstrate the integration of sustainability into academic research. This trend is further echoed in the academic output and citation impact of these regions.

In contrast, the SDGI and EPI rankings of the three countries of the United States, China, and India are far from their academic performance rankings. Despite having unparalleled academic results in their financial research, they are not ranked among the best when measured against other institutions for performance. This suggests that there is room for greater integration of environmental performance within its finance scholarship. This contrast serves as a stark reminder of the potential disconnect between research quantity and applied sustainability outcomes.

[PLEASE INSERT Figure 6 and Figure 7 HERE]

As shown in Figure 6 and Figure 7, the two stacked bar charts provide a comprehensive view of the financial journal publications and citation trend related to the UN SDGs from 2010 to 2022 emerges, highlighting contributions from the top five countries. Looking at the overall trend, the number of publications is probably on the rise, especially since 2015, which shows that scholars are paying more and more attention to this field.

The United States(the US) has maintained a steady lead in publication volume and citation percentage over the observed period, emphasizing its dominant role in advancing research in this area. However, the growth in China's both publication count and citation ratio particularly notable. Beginning from a low base of publications in 2010, we see a consistent increase, becoming a significant contributor by 2022. After 2017, it overtook the UK and became a stable second place in publication area. This shows that while China is paying more and more attention to the sustainable side and is committed to expanding its research , the current focus may be more on expanding numbers rather than conducting research with universal impact, although the impact is gradually increasing. Conversely, the United Kingdom(the UK), Germany, and Australia, though consistently contributing, maintained a relatively steady publication rate, reflecting a stable research environment.

11

In the geographical context, academic performance and sustainable develop performance evaluation data reveal a divergence in research strategies: China is actively expanding its academic influence, yet its performance in sustainable development has not entirely kept pace with its scholarly advancements. Conversely, Germany, France, Italy, and Spain demonstrate consistent leadership in both academic and practical realms of sustainable development. Notably, there exists a significant gap between India's sustainable development performance and its academic impact. This divergence offers critical insights into the correlation between academic contributions and actual sustainable outcomes, highlighting the varied effectiveness of research in driving practical advancements.

4.1.3. Journal Publication analysis

[PLEASE INSERT Table 5 HERE]

[PLEASE INSERT Figure 8 and Figure 9 HERE]

Figure 8 and Figure 9, along with Table 5 shows the top ten SDG-related journals. Especially starting from 2020, the number of financial papers published related to UN SDGs has surged, reflecting the increasing focus on sustainability in the academic community. This surge is likely linked to the increasing momentum towards the 2030 Agenda for Sustainable Development.

The annual publication data indicates that finance journals are incrementally shifting their focus towards SDGs. We can observe two striking polar opposites in terms of experience. Examining the proportion of papers in the top 10 journals by journal we can see which are gaining and which losing market share of SDG publications. One journal 'Finance Research Letters' in particular has shown a significant increase in SDG-related articles since 2017 signaling the journal's strategic orientation towards integrating sustainability in finance. This trend is further substantiated by a rise in citation counts from 2017, suggesting that the research community is producing not only more content on UN SDGs but also work that is gaining increasing recognition and influence in the academic field. On the contrary, it is very obvious that the number of SDG related publications in the 'Journal of Banking and Finance' during the past twelve years has basically remained unchanged, but the proportion of citations has decreased year by year. This shows that the journal's influence in this area may be declining year by year.

12

These trends highlight the growing incorporation of sustainable development goals into the core of financial research, with leading journals acknowledging and enhancing the significance of research that aligns with global sustainability objectives. This evolution reflects the dynamic nature of finance as a discipline, which is increasingly addressing the critical environmental and social challenges encapsulated in the UN SDGs.

4.1.4. ABS Level Analysis

This section focuses on examining the performance of Academic Journal Guide 2021, also known as ABS journals at different levels. This study uses the financial field of ABS2021 as the basis for classification.

[PLEASE INSERT Table 6 and Table 7 HERE]

[PLEASE INSERT Figure 10 and Figure 12 HERE]

The data in Table 6 and Table 7 are statistics and percentage charts on the number of publications, based on ABS and 17 UN SDGs. As shown in Figure 10 and Figure 12, analyzing publication trend and citation trend in finance journals from 2010 to 2022, categorized by ABS rankings and aligned with UN SDGs, the study reveals distinct thematic priorities and journal quality characteristics. Notably, themes such as Economic Growth(Goal 8) and Peace and Justice(Goal 16) are predominant in both publication and citation area, reflecting their wider impact and urgency on a global scale. This suggests a strong academic focus in these areas. In contrast, topics like Oceans(Goal 14), Land Life(Goal 15), and Water and Sanitation(Goal 6) have garnered less attention, indicating potential research gaps or challenges in publication within these fields. The bar chart clearly illustrates how journal rankings correlate with thematic focus. The chart presents that ABS 3 journals, which are in the mid-tier category, are the most prominent. They lead in terms of both publication and citation volumes across various fields related to the UN SDGs.

This shows that these journals are actively involved in disseminating various research related to sustainable development and have great influence. In contrast, the highest-ranked ABS 4* journals have expectantly low publication numbers but as we would expect have very high citations. Somewhat surprisingly, ABS4 journals did not perform well. Their number of publications was basically the same as ABS4*, but their number of citations was low. This may reflect that the ABS 4 journal's selection of topics is relatively conservative.

13

As a collection of highly rated journals, its influence is not as influential as the ABS3 journal collection.

Across the areas there is broad similarity as to the distribution of papers by SDG across the ABS categories. The percentage of papers in ABS2 is generally similar to that of ABS4*. To empirically test our conjecture regarding the uniformity of SDG research distribution across different ABS categories a chi-square test was conducted.

The formulated hypotheses for this analysis were as follows:

Null Hypothesis (H0): The distribution of SDG papers across different ABS categories is uniform, with no significant differences.

Alternative Hypothesis (H1): The distribution of SDG papers across different ABS categories is not uniform, exhibiting significant differences.

[PLEASE INSERT Table 8 HERE]

The chi-square statistic for each category, as detailed in the Table 8, was calculated to ascertain the distribution pattern. Upon computation, the chi-square test yielded a cumulative statistic of 638.34. Set against a critical value of 93.22 at a significance level of 0.01 with 64 degrees of freedom, this statistic markedly surpasses the threshold. This result decisively leads us to reject the null hypothesis, indicating that the distribution of papers is significantly non-uniform across different ABS categories. Such a finding suggests the presence of a disparity in how research themes aligned with UN SDGs are represented in various academic tiers as classified by ABS. However, this non-uniform distribution does not preclude thematic overlap. While the volume of research in each ABS category varies, certain topics may recur across categories. This suggests that, despite the disparate quantity of research, there is a convergence around specific themes within sustainable development, indicative of shared scholarly focus. Concurrently, the presence of underrepresented topics across categories could signify potential areas for future research, aiming to address these gaps within the landscape of sustainable development studies.

There are however some issues. For those areas with more than 5% of the total overall (Goal 8 Economic Growth, Goal 10 Inequality, Goal 12 Sustainable Consumption and Goal 16 Peace and Justice), we see some interesting patterns.

The UN SDGs on Goal 10 (Inequality), Goal 12 (Sustainable Consumption) and Goal 16 (Peace and Justice) show an a falloff from ABS1 to ABS 4*. In other words the ABS

14

4* journals are less interested than the ABS 1 journals in these areas, proportionally. For these the fall is monotonic across areas, increasing in each cohort. In Goal 8(Economic Growth), we see that it is high in the middle and low on both sides, which means that ABS 3 is the most concerned about this topic, and ABS 1 and ABS 4* are both less concerned than ABS.

The study also identifies potential research opportunities and existing hotspots. Fewer publications in areas like Oceans and Land Life, suggest unexplored research territories, while the high volume of work on Economic Growth and Inequality indicates well-established areas of academic inquiry.

This comprehensive analysis provides a nuanced understanding of the academic landscape regarding the UN SDGs. It aids researchers and policymakers in identifying focus areas for future investigation and resource allocation, shaping the trajectory of research and policy in finance-related fields.

4.2. Topic Analysis

In this subsection, we apply Latent Dirichlet Allocation (LDA) Topic Modelling. Initially, to ascertain the number of LDA topics to investigate, we calculated the topic perplexity.

The Figure 13 elucidates the trend in perplexity relative to the number of topics in the LDA model, demonstrating a marked decrease in perplexity up to $n=6$, after which the rate of decline moderates, indicating a point of diminishing returns in terms of model complexity. Although the perplexity at $n=6$ is not the absolute lowest, it represents an optimal balance, making it a suitable choice for the number of topics. In the figure, $n=12$ is the lowest point of topic perplexity, this is an excessive number of topics, which may lead to overfitting. Therefore, based on this graphical analysis, we selected six topics as a reasonable starting point that offers a good equilibrium between model complexity and interpretability.

Following the assessment of the optimal number of topics through perplexity analysis, we proceeded with the Topic Similarity Matrix visualization. As shown in Figure 14, the heatmap constructed on Jensen-Shannon divergence values within a six-topic LDA model, provides a measure of distinction between the topics. The diagonal entries naturally approach zero divergence, as evidenced by the lighter shades, confirming the identity of topics with themselves. Conversely, the off-diagonal entries display predominantly darker shades,

15

indicating that the topics are well-differentiated, each encapsulating unique facets of the dataset. Noteworthy observations include the cells between topics 3 and 5 and topics 4 and 6, which are lighter, suggesting a higher degree of similarity and possible thematic intersections. These particular similarities may point to shared keywords or concepts between these topic pairs.

[PLEASE INSERT Figure 13 and Figure 14 HERE]

[PLEASE INSERT Table 8 HERE]

This analysis, leveraging abstracts from financial journals, uncovers a multifaceted interface with various UN SDGs, reflecting the expansive social and environmental repercussions of financial activities. The themes identified through the LDA examination of financial texts related to the UN SDGs shed light on focal points within the realm of sustainable finance. Keywords such as 'risk', 'insurance', and 'financial value' in Topic 1 underscore an acute awareness of valuation in risk management and the integrity of financial markets. Topics 2 and 3, encompassing terms like 'market', 'stock', 'crisis', 'rate', and 'policy', signal scholarly attention on market behaviors, crisis response, interest rate fluctuations, and economic policy impacts. The prominence of 'firms', 'banks', 'csr', and 'corporate' within Topic 4, along with 'performance', 'governance', and 'board' in Topic 5, indicates an analytical focus on the internal mechanisms of corporate entities and the strategic oversight of operational and financial risks. Finally, Topic 6, with terms like 'financial growth', 'economic development', 'green', and 'energy', accentuates the sector's engagement with economic growth strategies and sustainable development, highlighting an integrative approach to finance that encompasses both profitability and social responsibility. These topics collectively point to a robust discourse recognizing the integral role of financial sectors in addressing global economic challenges and sustainability.

[PLEASE INSERT Figure 15 and Figure 16 HERE]

Building upon the thematic exploration of financial journal abstracts in relation to the UN SDGs, further analysis reveals correlations between specific goals and the thematic concentration of scholarly articles. Notably, the majority of papers under Goals 1(Poverty) and 12(Sustainable Consumption) predominantly address 'Economic policy', highlighting a scholarly predilection for policy discourse within these areas. Goal 5's(Gender Equality)

16

most salient theme is 'Corporate Finance', indicating that current research on gender equality is heavily focused on corporate finance sectors, potentially exploring directions such as gender diversity in financial leadership or the impact of financial policies on gender equity.

Moreover, Goals 7(Energy), 8(Economic Growth), 10(Inequality), and 17(Partnerships) share a substantial focus on 'Market dynamics', underscoring the sector's keen interest in the fluctuations and intricacies of market systems. Both Goals 7(Energy) and 12(Sustainable Consumption) show a significant portion of research dedicated to 'Green Economy', reflecting the strong interrelation between sustainable economic practices, energy, and green consumerism. Goal 8(Economic Growth), which boasts the highest volume of publications, extensively discusses 'Market Dynamics' and 'Corporate Governance', suggesting that the intricacies of market behavior and the structures of corporate oversight are pivotal concerns within this SDG.

In a broader analysis, topics under 'Market Dynamics' are most prevalent across the UN SDGs corpus, indicating an overarching interest in the forces that drive financial markets. This is closely followed by research in 'Corporate Finance', signaling a robust discourse on the financial decision-making within firms. Contrastingly, 'Corporate Governance' receives relatively minimal focus, which may suggest a research gap in the literature. This gap could be indicative of an opportunity for future studies to delve deeper into the frameworks and practices that govern corporate entities, particularly in light of Goal 8's (Economic Growth) dominance, which may disproportionately influence the overall thematic distribution observed in the dataset.

5. Conclusions

This research conducts an empirical analysis of the correlation between the UN Sustainable Development Goals (SDGs) and publications in leading financial journals over the past twelve years. The study employs a Fuzzy Match algorithm, enhanced by the Rake algorithm for feature extraction, to improve classification accuracy, particularly with complex datasets. The classifier model used achieves an accuracy of approximately 82

The findings highlight a significant alignment of finance research with specific UN SDGs. However, areas like marine ecology, water management, and terrestrial life are less represented, suggesting opportunities for further research. Analysis of publication and citation trends reveals a varied landscape, with countries exhibiting either a broad research spec-

17

trum or focusing on influential works with widespread recognition. Over time, there's a growing shift in scholarly focus towards sustainable development.

Geographically, the United States and China are consistently leading in terms of publication and citation numbers, with the United Kingdom ranking third. Notably, the United States, despite its academic leadership in financial research, does not rank as high in the Environmental Performance Index (EPI) and the Sustainable Development Goals Index (SDGI), suggesting a disparity between its environmental performance and academic influence. Similarly, the practical sustainability performance of China and India falls short of their academic achievements. The research also compares academic rankings with sustainable development rankings, finding that the academic performance of Germany, Spain, the UK, and France aligns more closely with their performance in sustainable development.

The study examines publication trends in journals ranked by the Academic Journal Guide (ABS) in relation to the United Nations Sustainable Development Goals (UN SDGs), uncovering a varied academic landscape. It observes a notable contribution of mid-tier ABS 3 journals in circulating sustainability-related research. In contrast, the highest-ranking ABS 4* journals, despite having fewer publications, exhibit a higher impact based on citations. Meanwhile, ABS 4 journals show a lower performance, with minimal publications and citations, differing significantly from both ABS 4* and ABS 3 journals. The analysis also highlights discrepancies in the representation of UN SDG-related research themes across different ABS journal categories. It uncovers a non-uniform distribution of research across these categories, while also identifying common thematic areas that indicate a collective academic interest and opportunities for further research in sustainable development. The application of an LDA Topic Modeling technique in this analysis surfaces a shift in financial research from traditional themes to those that include global socio-economic and environmental goals. The identified topics - 'Financial Risk and Valuation', 'Market Dynamics and Volatility', 'Economic Policy and Rates', 'Corporate Finance and Banking', 'Corporate Governance and Performance', and 'Green Economics' - demonstrate an expanded focus on sustainability within scholarly work. The prevalence of 'Market Dynamics' in the discourse suggests a strong interest in the changing dynamics of financial markets, while the lesser focus on 'Corporate Governance' may highlight an area with potential for further detailed research. These trends in financial research literature reflect its evolving role in addressing key global challenges and a shift towards sustainability and responsible governance in the financial sector.

In summary, this study examines the intersection of financial research and global sustainable development goals. It provides insights that integrate core principles of finance with the evolving global conversation on sustainability. This ongoing exploration highlights the relationship between finance and sustainability, outlining potential directions for future research and influencing policy development. The aim is to guide financial strategies towards sustainability and resilience, while considering societal needs. The findings of this study have implications for various groups, including academics and policymakers, underlining the continued relevance and influence of financial research in the contemporary context.

Bibliography

References

Ahsan, M. M., F. Z. Chowdhury, M. Sabilah, A. W. B. A. Wahab, and M. Y. I. B. Idris (2017). An efficient fuzzy keyword matching technique for searching through encrypted cloud data. In *2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*, pp. 1–5. IEEE.

Alcántara-Rubio, L., R. Valderrama-Hernández, C. Solís-Espallargas, and J. Ruiz-Morales (2022). The implementation of the sdgs in universities: A systematic review. *Environmental Education Research 28*(11), 1585–1615.

Aria, M. and C. Cuccurullo (2017). bibliometrix: An r-tool for comprehensive science mapping analysis. *Journal of informetrics 11*(4), 959–975.

Atz, U., T. Van Holt, Z. Z. Liu, and C. C. Bruno (2023). Does sustainability generate better financial performance? review, meta-analysis, and propositions. *Journal of Sustainable Finance & Investment 13*(1), 802–825.

Ball, R. (2017). *An introduction to bibliometrics: New development and trends.* Chandos Publishing.

Bebbington, J. and J. Unerman (2018). Achieving the united nations sustainable development goals: an enabling role for accounting research. *Accounting, Auditing & Accountability Journal 31*(1), 2–24.

Beck, T., A. Demirgüç-Kunt, and R. Levine (2006). Bank supervision and corruption in lending. *Journal of monetary Economics 53*(8), 2131–2163.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent dirichlet allocation. *Journal of machine Learning research 3*(Jan), 993–1022.

Cayrol, M., H. Farreny, and H. Prade (1982). Fuzzy pattern matching. *Kybernetes 11*(2), 103–116.

Chaudhuri, S., K. Ganjam, V. Ganti, and R. Motwani (2003). Robust and efficient fuzzy match for online data cleaning. In *Proceedings of the 2003 ACM SIGMOD international conference on Management of data*, pp. 313–324.

Cheng, I.-H., H. Hong, and K. Shue (2023). Do managers do good with other people's money? *The Review of Corporate Finance Studies 12*(3), 443–487.

de Souza Barbosa, A., M. C. B. C. da Silva, L. B. da Silva, S. N. Morioka, and V. F. de Souza (2023). Integration of environmental, social, and governance (esg) criteria: their impacts on corporate sustainability performance. *Humanities and Social Sciences Communications 10*(1), 1–18.

Donthu, N., S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of business research 133*, 285–296.

Edmans, A. (2023). The end of esg. *Financial Management 52*(1), 3–17.

Flammer, C., T. Giroux, and G. Heal (2023). Biodiversity finance. Technical report, National Bureau of Economic Research.

Karolyi, G. A. and J. Tobin-de la Puente (2023). Biodiversity finance: A call for research into financing nature. *Financial Management*.

Khan, A., J. W. Goodell, M. K. Hassan, and A. Paltrinieri (2022). A bibliometric review of finance bibliometric papers. *Finance Research Letters 47*, 102520.

Kumar, S., D. Sharma, S. Rao, W. M. Lim, and S. K. Mangla (2022). Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research. *Annals of Operations Research*, 1–44.

Li, G., X. Zhu, J. Wang, D. Wu, and J. Li (2017). Using lda model to quantify and visualize textual financial stability report. *Procedia computer science 122*, 370–376.

Lucey, B. M., S. Kumar, and R. Sureka (2023). Corruption in finance research: The state of art and future research agenda. *Journal of Economic Criminology*, 100001.

Mbarki, I., M. A. Khan, S. Karim, A. Paltrinieri, and B. M. Lucey (2023). Unveiling commodities-financial markets intersections from a bibliometric perspective. *Resources Policy 83*, 103635.

Oestreich, J. E. (2023). Closing the circle of implementation: the sustainable development goals, universal periodic review, and the rights-based approach to development. *The International Journal of Human Rights*, 1–20.

Pandey, D. K., B. M. Lucey, and S. Kumar (2023). Border disputes, conflicts, war, and financial markets research: A systematic review. *Research in International Business and Finance*, 101972.

Renneboog, L., J. Ter Horst, and C. Zhang (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of banking & finance 32*(9), 1723–1742.

Sachs, J. D. (2012). From millennium development goals to sustainable development goals. *The lancet 379*(9832), 2206–2211.

Song, X., G. Li, J. Feng, and L. Zhou (2009). Effective fuzzy keyword search over uncertain data. In *Database Systems for Advanced Applications: 14th International Conference, DASFAA 2009, Brisbane, Australia, April 21-23, 2009. Proceedings 14*, pp. 66–70. Springer.

United Nations (2015). Transforming our world: the 2030 agenda for sustainable development. Available at: https://undocs.org/en/A/RES/70/1. Accessed: 2023-12-01.

Wang, B., S. Yu, W. Lou, and Y. T. Hou (2014). Privacy-preserving multi-keyword fuzzy search over encrypted data in the cloud. In *IEEE INFOCOM 2014-IEEE conference on computer communications*, pp. 2112–2120. IEEE.
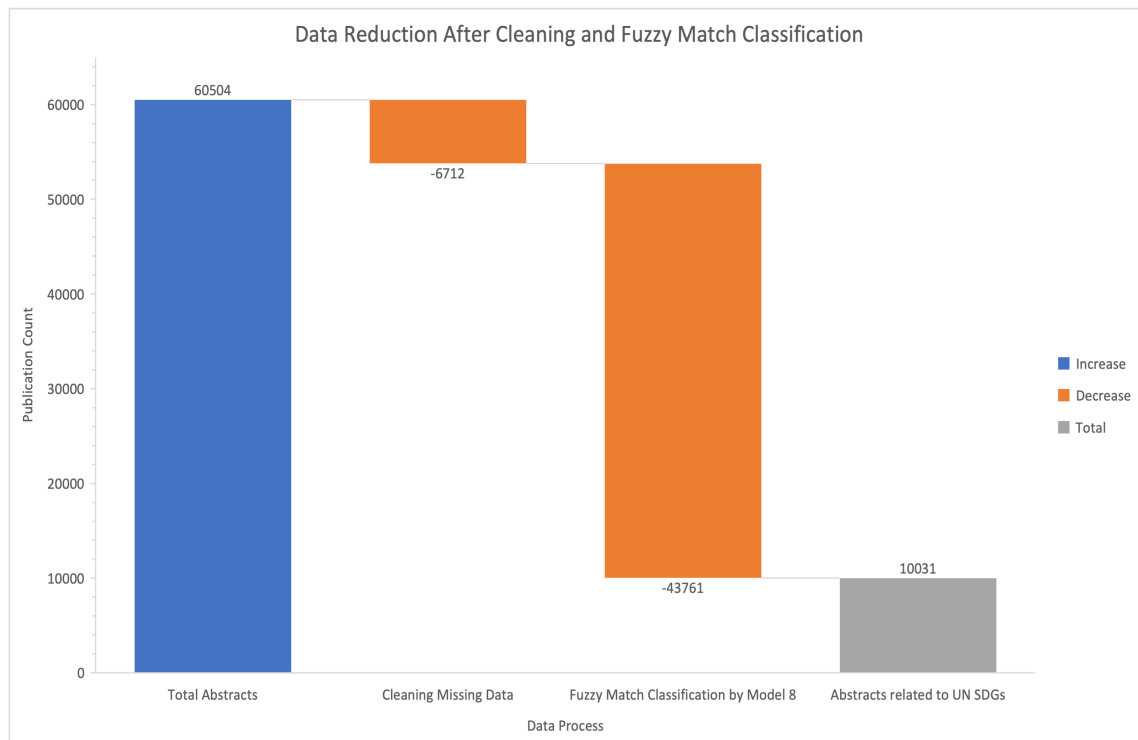
Appendix

Appendix A - Figures

Figure 1: Data Reduction After Cleaning and Fuzzy Match Classification
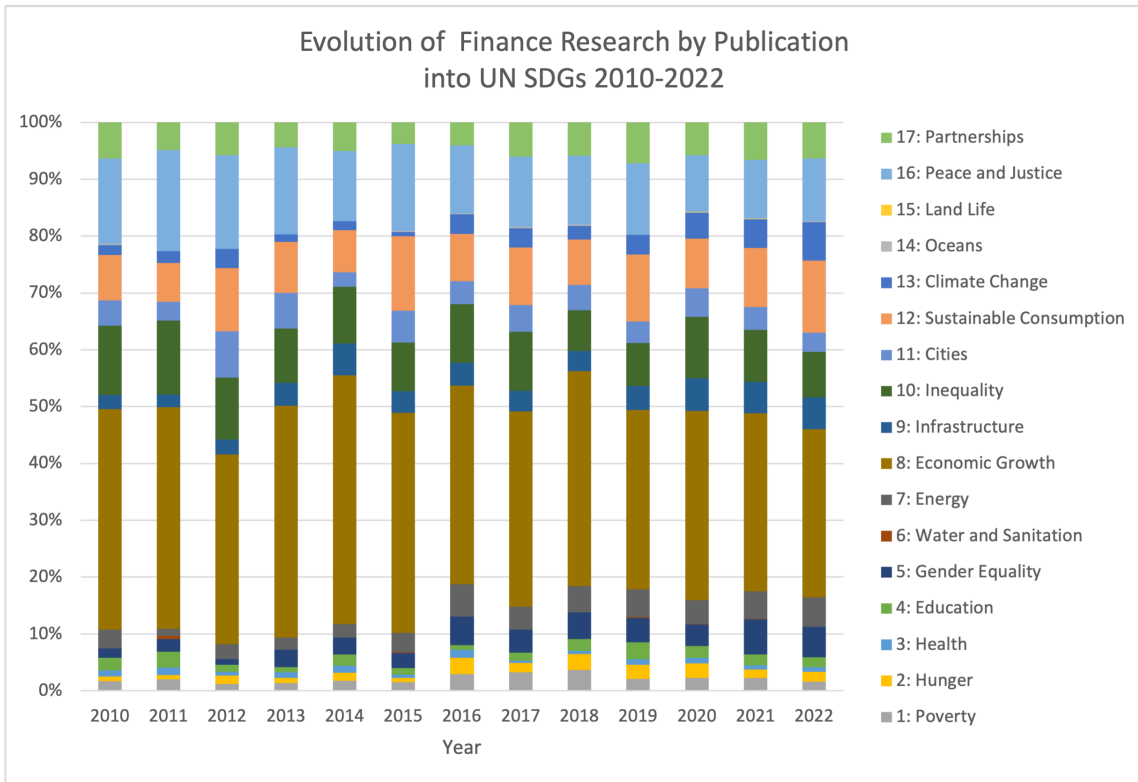
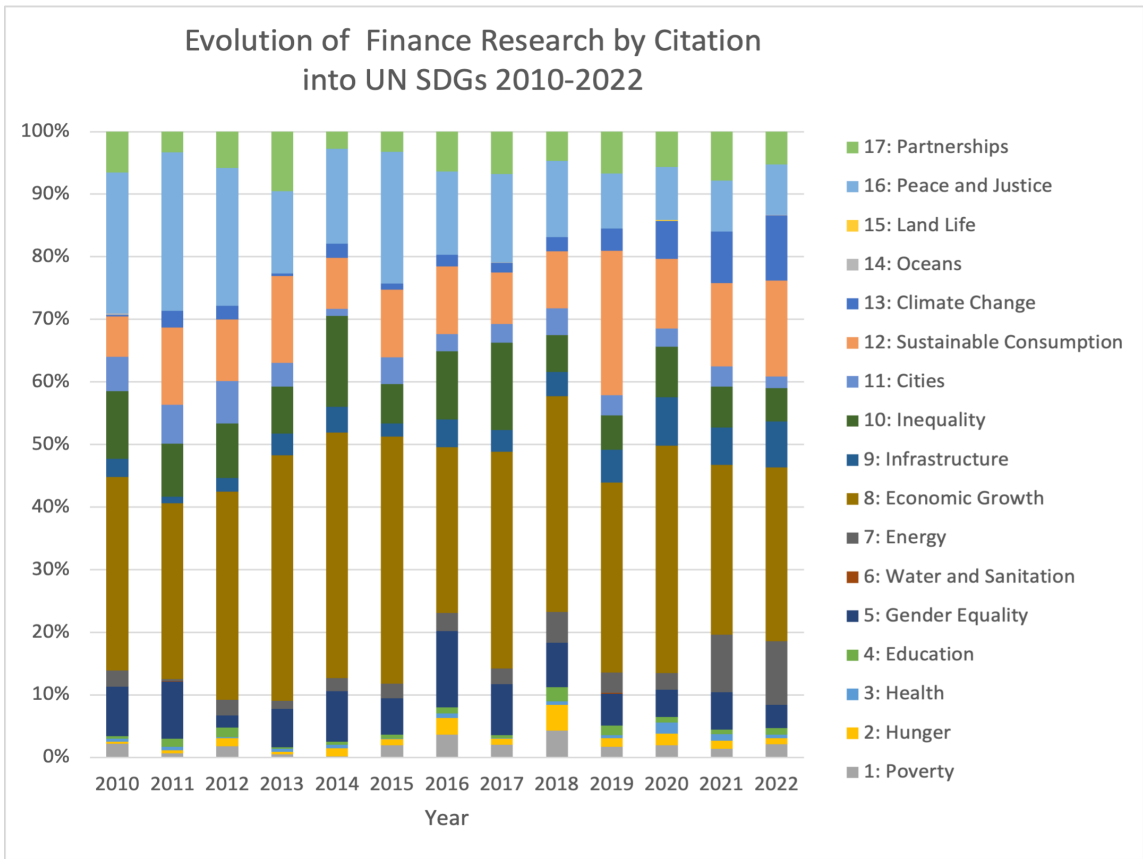Figure 2: Evolution of Finance Research by Publication into UN SDGs 2010-2022

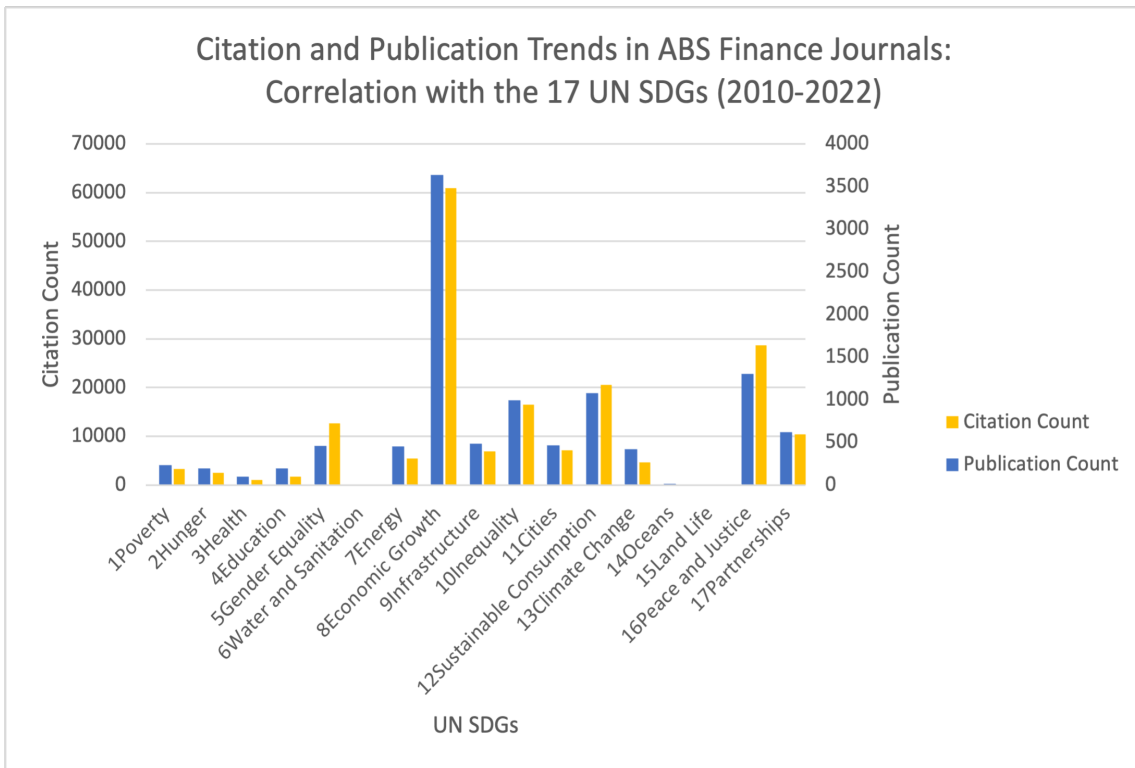Figure 3: Evolution of Finance Research by Citation into UN SDGs 2010-2022

Figure 4: Citation and Publication Trends in ABS Finance Journals:
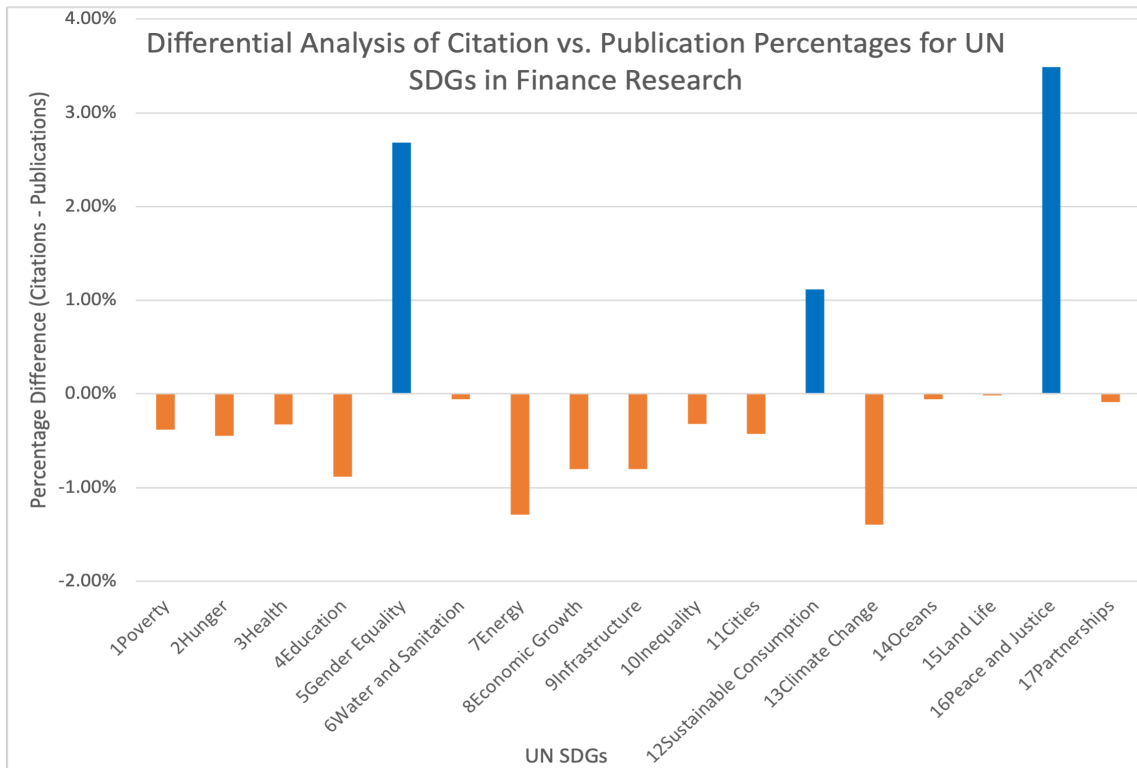Correlation with the 17 UN SDGs (2010-2022)

Figure 5: Differential Analysis of Citation vs. Publication Percentages for UN SDGs in Finance Researc
Notes: The bar chart depicts the percentage difference calculated by subtracting the publication percentage from the citation percentage for each UN SDGs in finance research. Bars in blue indicate a positive difference where the citation percentage exceeds the publication percentage. Conversely, orange bars represent a negative difference, where the citation percentage is lower than the publication percentage.
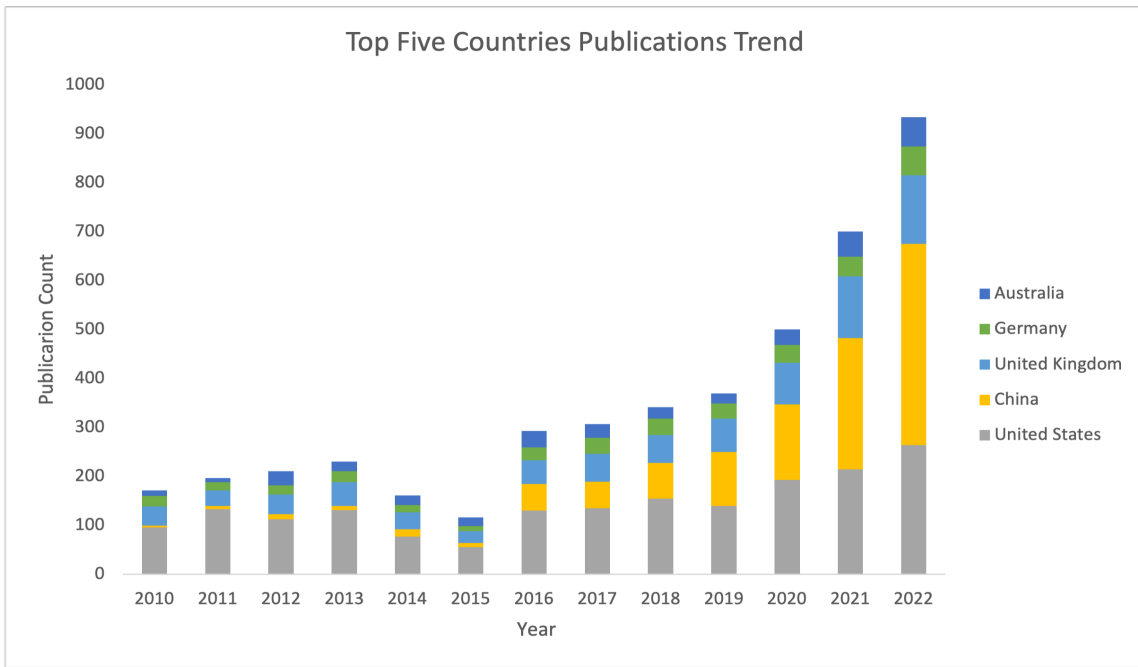
26
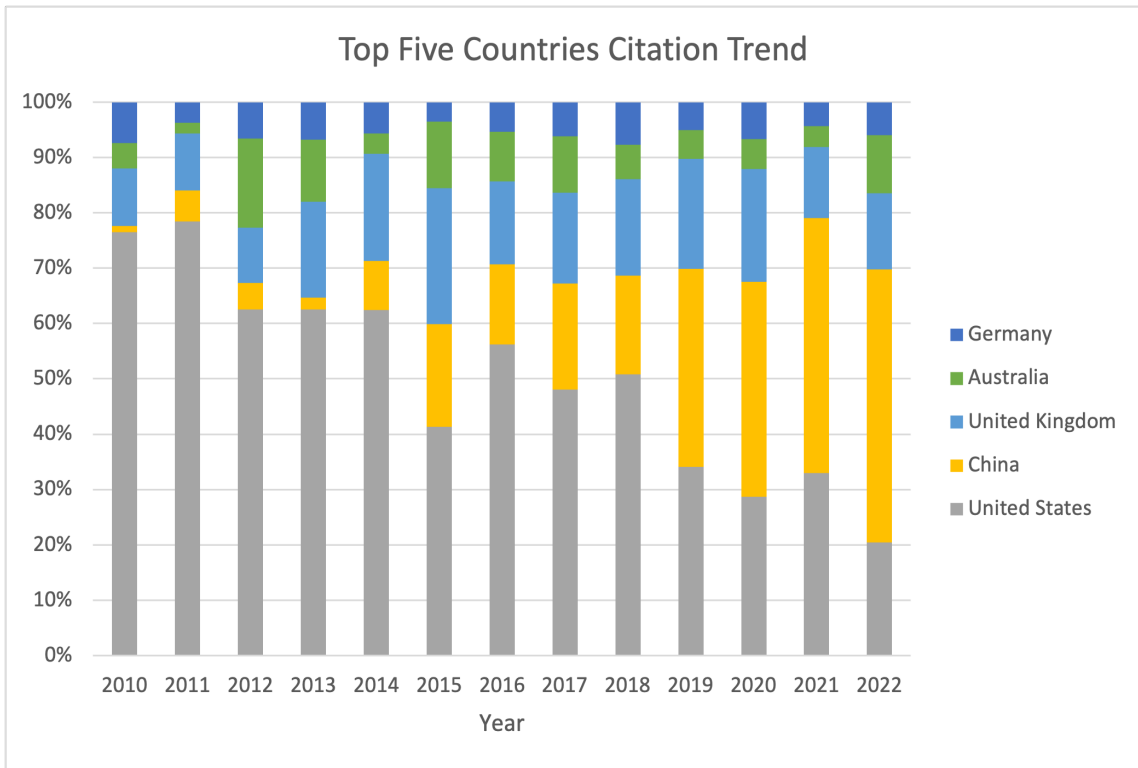
Figure 6: Top Five Countries Publication Trend
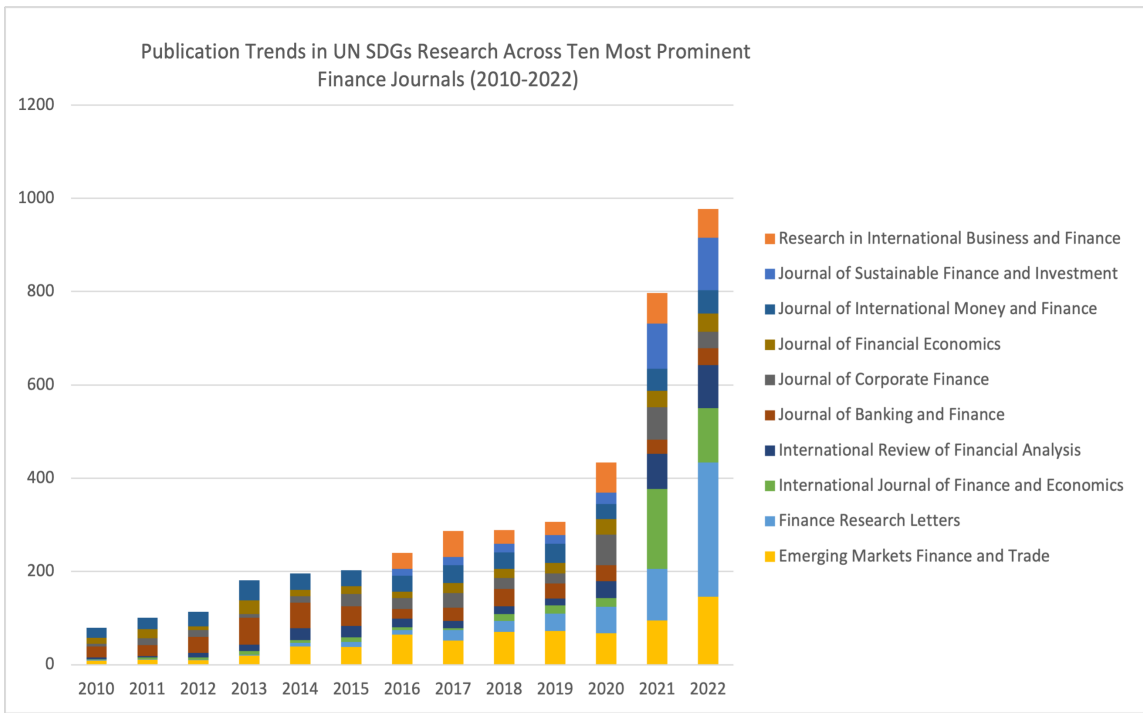
Figure 7: Top Five Countries Citation Trend

28

Figure 8: Publication Trends in UN SDGs Research Across Ten Most Prominent Finance Journals (2010-2022)
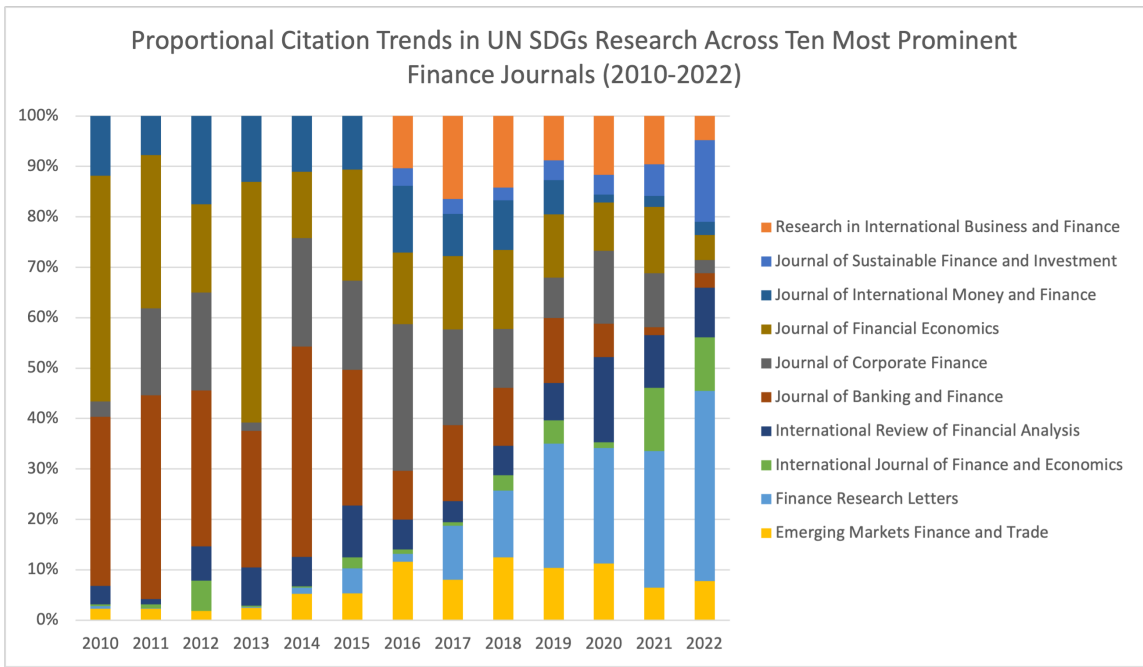
29

Figure 9: Proportional Citation Trends in UN SDGs Research Across Ten Most Prominent Finance Journals (2010-2022)
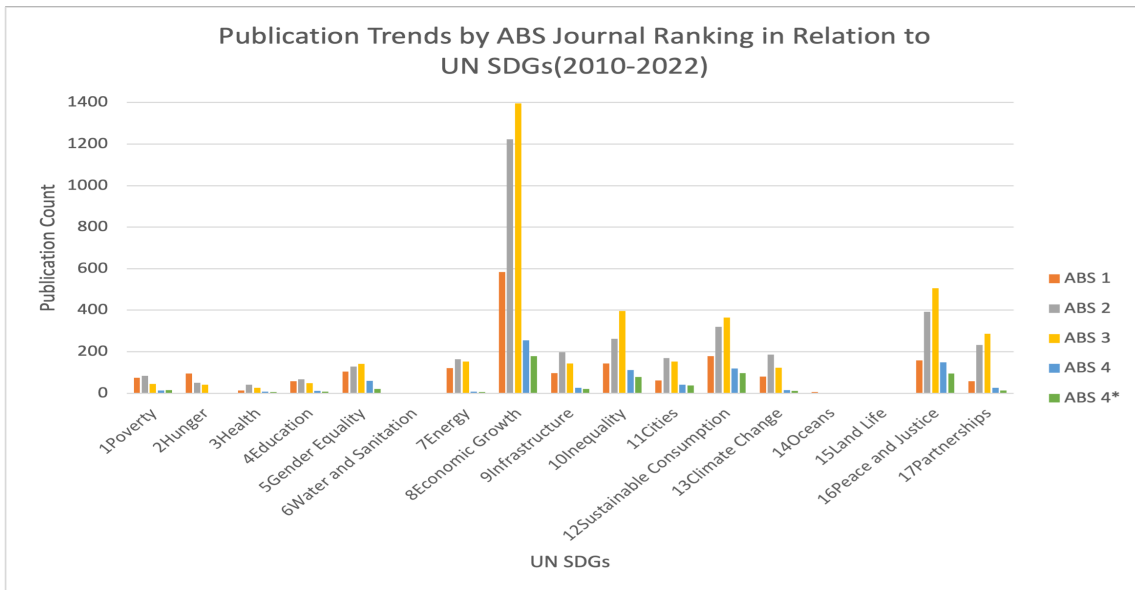
Figure 10: Publication Trends by ABS Journal Ranking in Relation to UN SDGs (2010-2022)

Electronic copy available at: https://ssrn.com/abstract=4696722

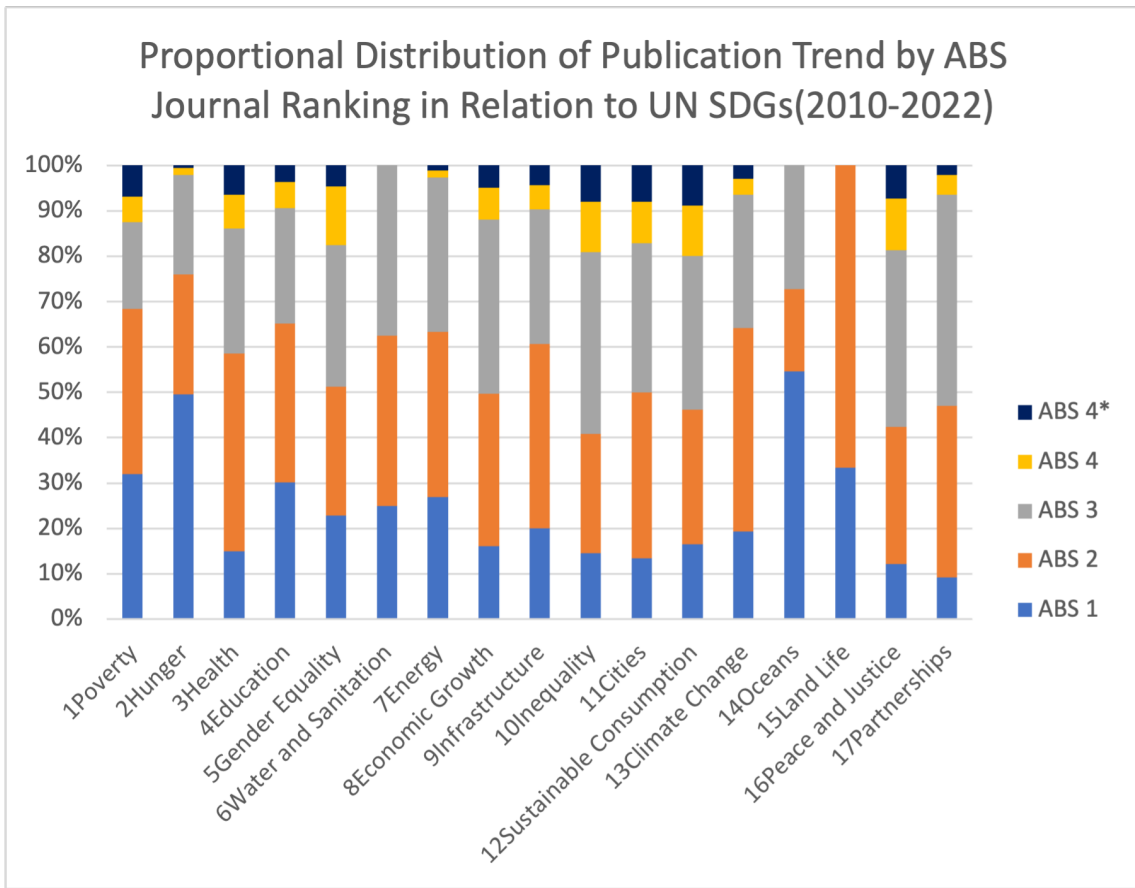Figure 11: Proportional Distribution of Publication Trend by ABS Journal Ranking in Relation to UN SDGs(2010-2022)
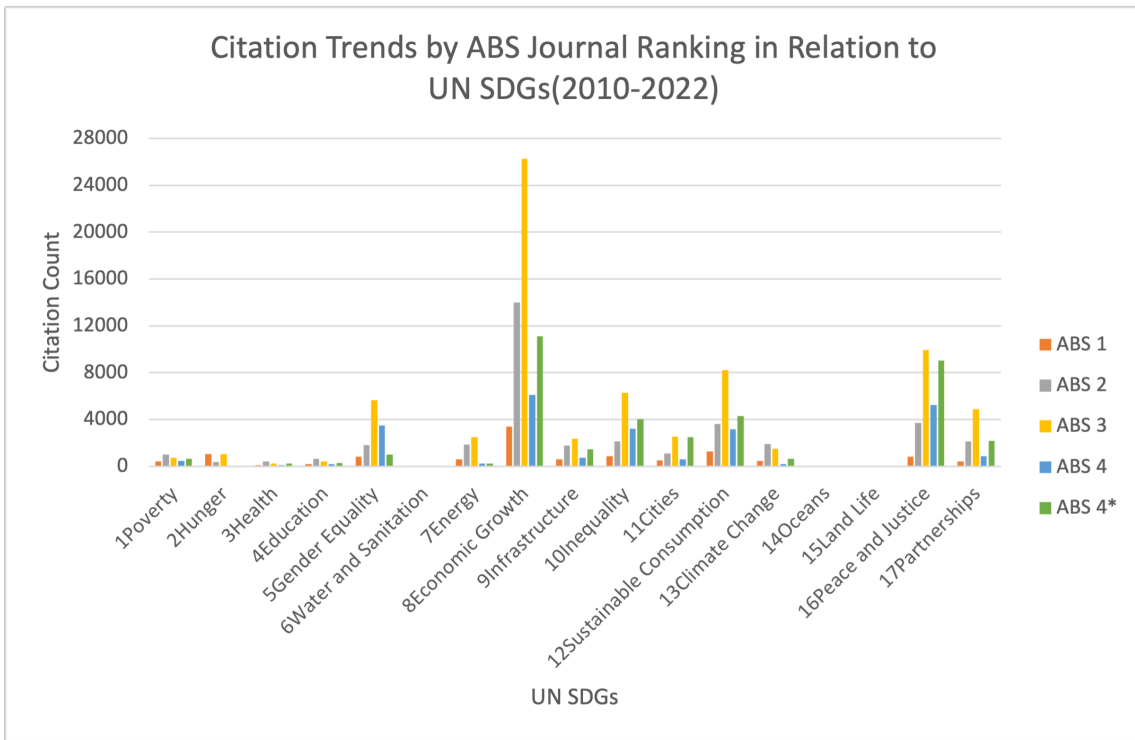
Figure 12: Citation Trends by ABS Journal Ranking in Relation to UN SDGs(2010-2022)

33
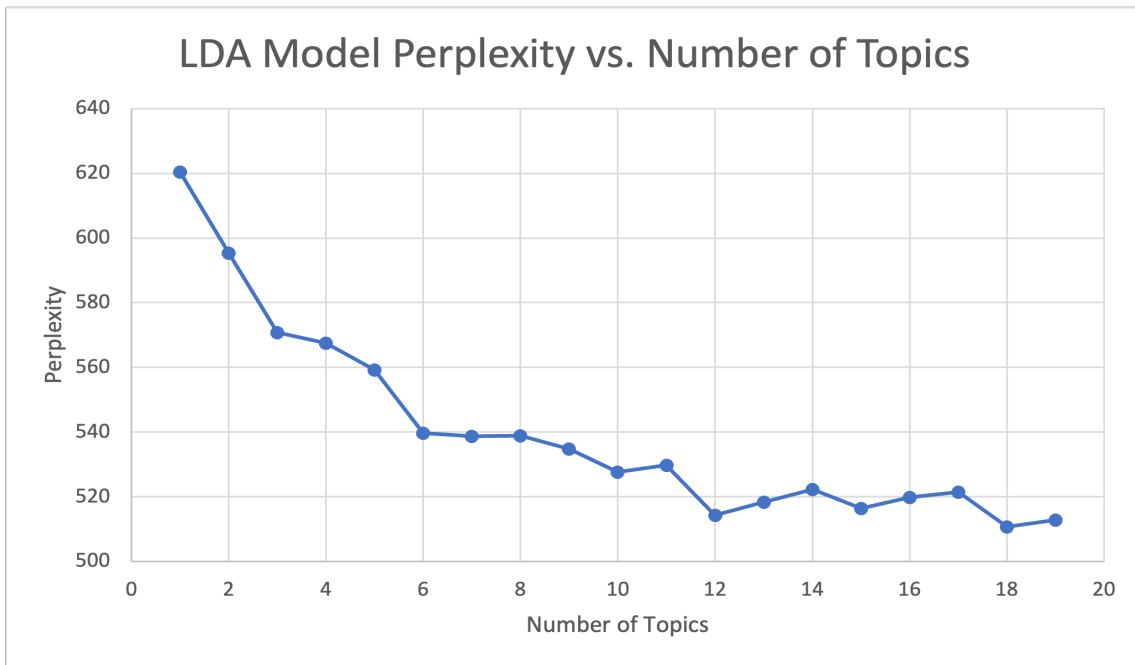
Figure 13: LDA Model Perplexity vs. Number of Topics

Perplexity measures the model's ability to predict new data, with lower values indicating better predictive performance. A steep decline in perplexity is observed as the number of topics increases, suggesting improved model fit. However, beyond a certain point, the benefits of adding more topics diminish, indicating a possible overfitting scenario where the model becomes too complex. In this chart, the leveling off of perplexity beyond 10 topics indicates a potential optimal range for topic selection.

34
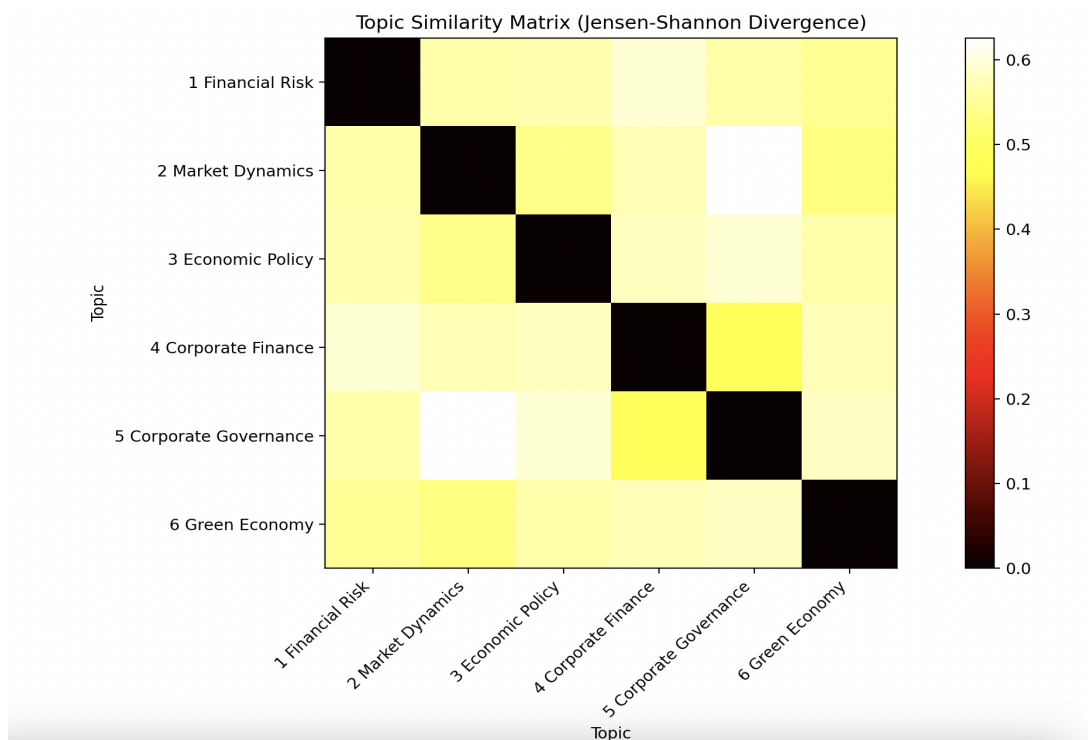
Figure 14: Topic Similarity Matrix (Jensen-Shannon Divergence)

Note: This heat map represents the Topic Similarity Matrix at n=6, utilizing Jensen-Shannon Divergence to quantify the similarity between topics generated by the LDA model. Darker squares indicate greater divergence or less similarity between topics, suggesting distinct thematic content, whereas lighter squares suggest higher similarity. The presence of darker squares along the matrix diagonal, where a topic is compared with itself, is expected and denotes zero divergence.

35
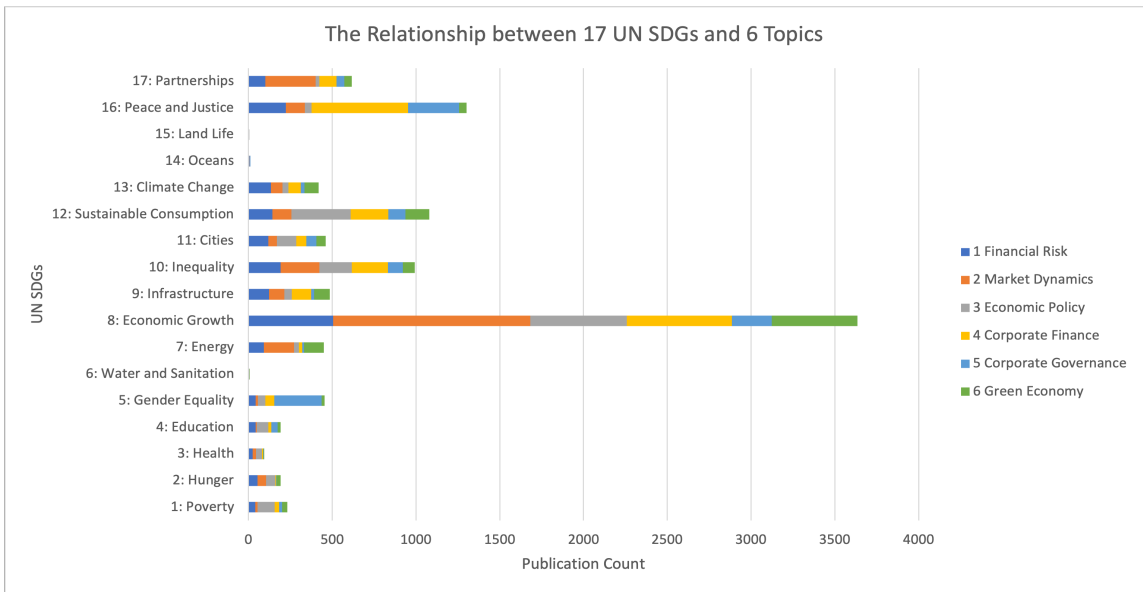
Figure 15: The Relationship between 17 UN SDGs and 6 Topics

36
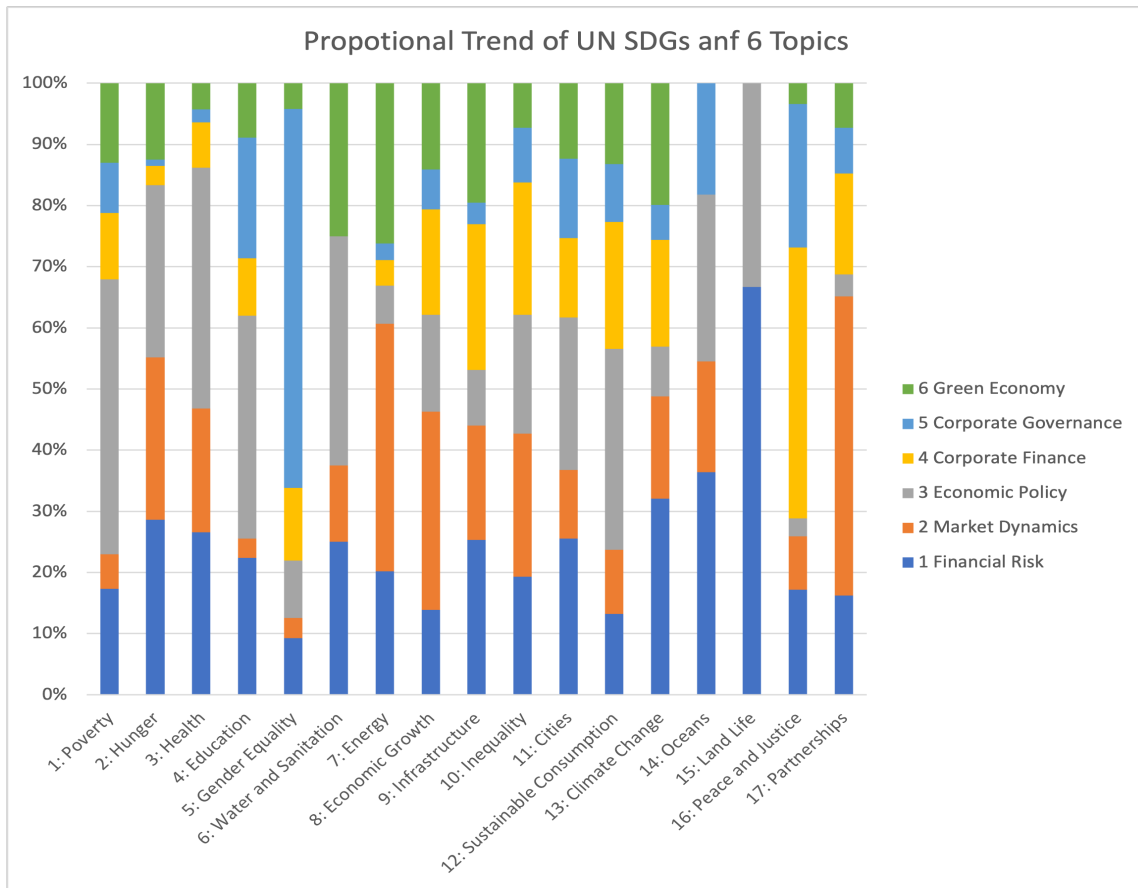
Figure 16: Proportional Trend of UN SDGs and 6 Topics

37
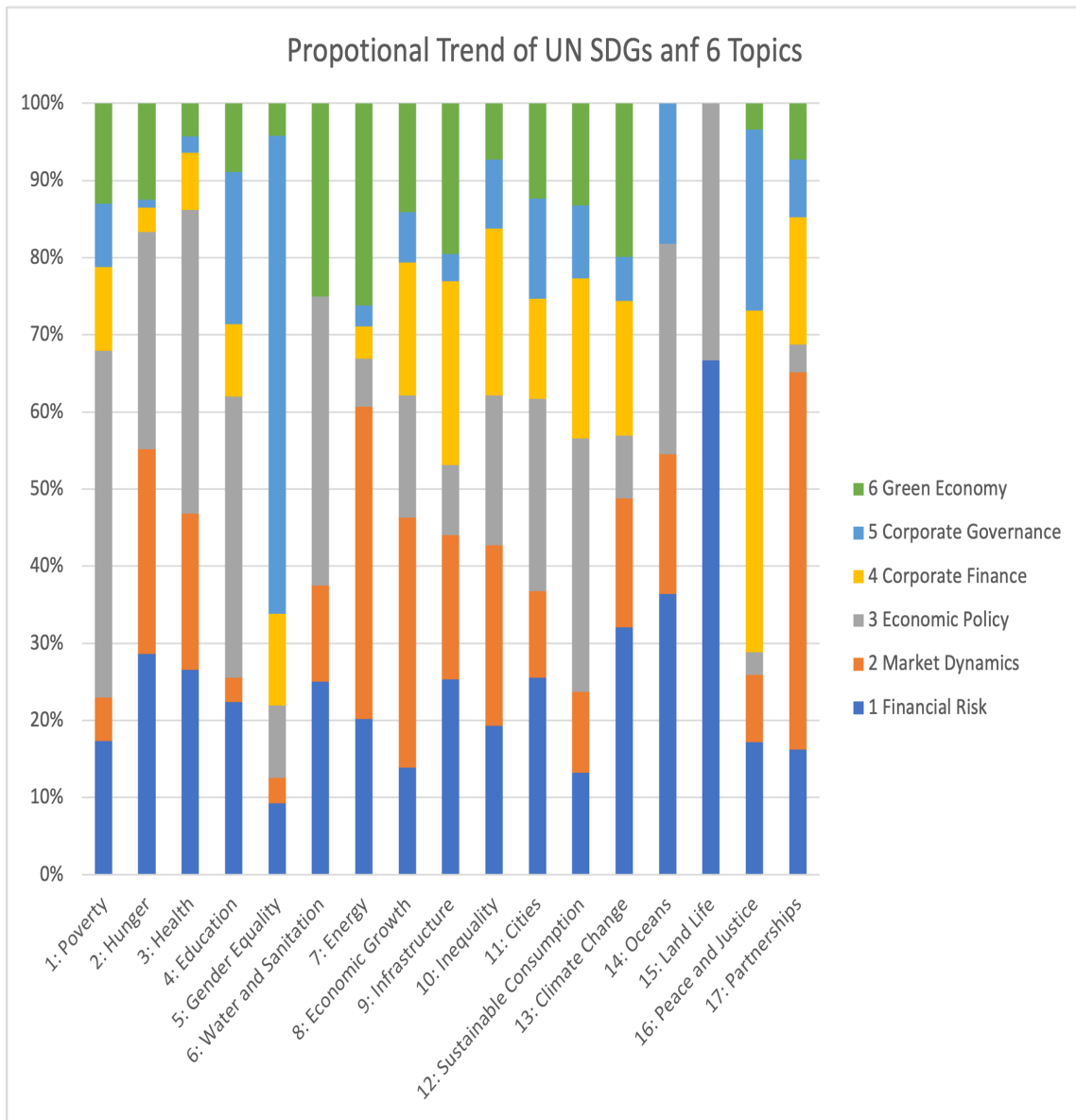
Figure 17: Distribution of SDG's across topics
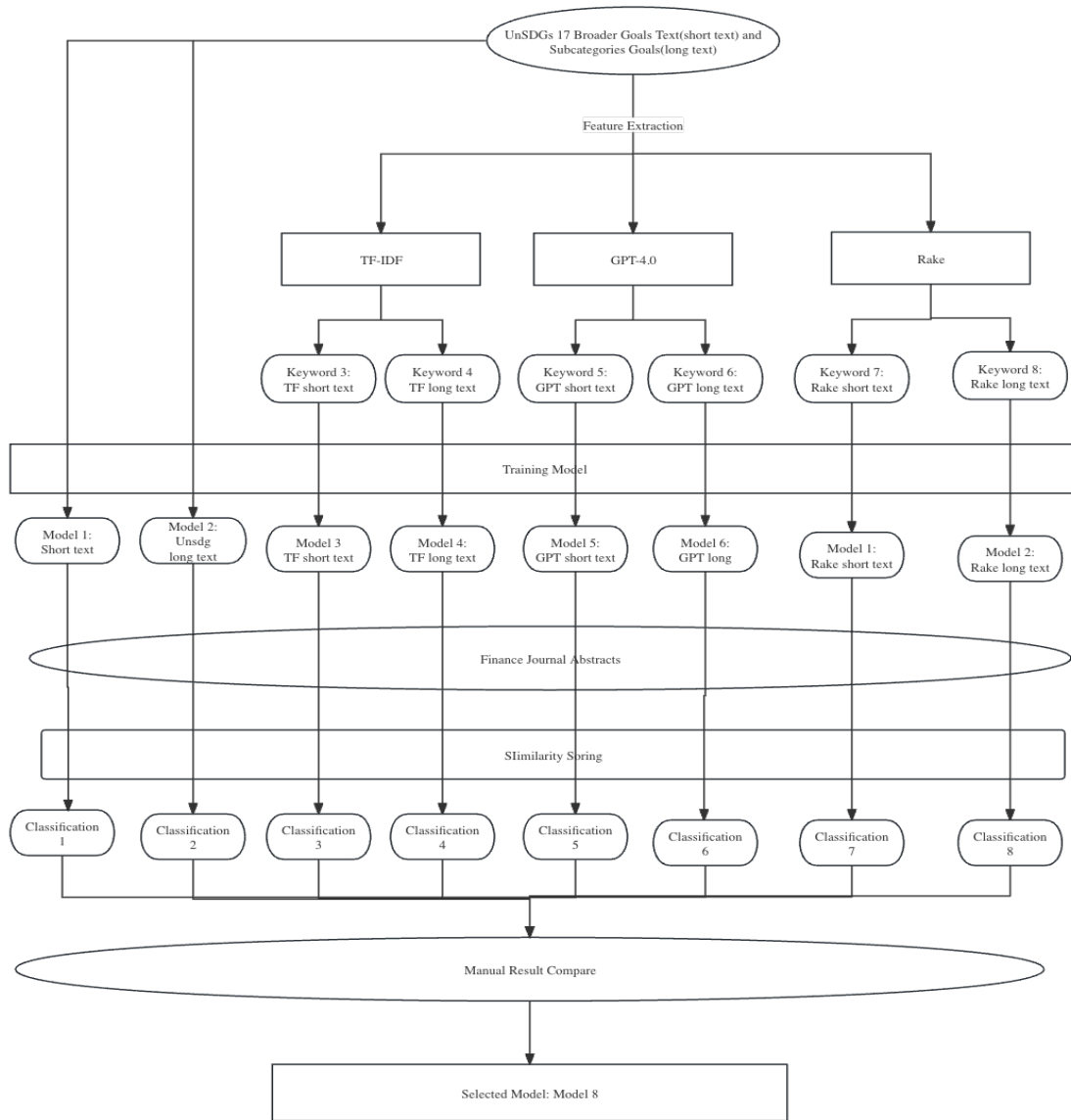
Figure 18: Fuzzy Match Application Process

39

Appendix B - Tables

Table 1: Model Names Table:

	UN SDGs broader goals text	**UN SDGs subcategories goals text**
Original Text	Model 1: Short text	Model 2: Long text
TF-IDF	Model 3: TF short text	Model 4: TF long text
GPT4.0	Model 5: GPT short text	Model 6: GPT long text
Rake	Model 7: Rake short text	Model 8: Rake long text

Notes: This table lists eight models for text analysis, organized along two dimensions: the focus of UN SDGs texts and the analysis method. The columns show models for texts targeting broader UN SDGs goals and those aimed at more specific subcategories. The rows display four types of models: Rake (Rapid Automatic Keyword Extraction), TF-IDF (Term Frequency-Inverse Document Frequency), GPT4.0, and Original Text, each applicable to either short or long text analysis.

Table 2: Accuracy Metrics

Model	Accuracy	False Positive Rate	False Negative Rate
Model 1: Short text	40.00%	41.25%	18.75%
Model 2: Long text	47.75%	35.00%	17.25%
Model 3: TF short text	52.00%	32.00%	16.00%
Model 4: TF long text	56.25%	16.00%	27.75%
Model 5: GPT short text	47.25%	5.25%	47.50%
Model 6: GPT long text	67.25%	26.50%	6.25%
Model 7: Rake short text	65.00%	32.25%	2.75%
Model 8: Rake Long text	80.75%	6.50%	12.75%

Notes: The "False Positives" and "False Negatives" represent two types of errors in model predictions. False Positives: Occur when a model incorrectly predicts a positive outcome. False Negatives: Happen when a model incorrectly predicts a negative outcome.

Accuracy = Number of Correct Predictions / Total Number of Predictions.

False Positive Rate = Number of False Positives / Total Number of Predictions.

False Negative Rate = Number of False Negatives / Total Number of Predictions.

Table 3: Citation and Publication Count in ABS Finance Journals

UN SDG	Publication Count	Citation Count
1 Poverty	204	2796
2 Hunger	166	2157
3 Health	90	948
4 Education	190	1579
5 Gender Equality	432	12268
6 Water and Sanitation	5	13
7 Energy	409	4949
8 Economic Growth	3416	58067
9 Infrastructure	473	6717
10 Inequality	903	14747
11 Cities	430	6681
12 Sustainable Consumption	1025	19691
13 Climate Change	405	4590
14 Oceans	8	25
15 Land Life	3	21
16 Peace and Justice	1275	28334
17 Partnerships	599	10123

Table 4: Publication Count and Environmental Rankings by Country

Country	Publication Count	Percentage of Publications	SDGI Rank	EPI Rank
United States	1838	18.32%	39	43
China	1178	11.74%	63	160
United Kingdom	799	7.97%	11	2
Germany	364	3.63%	4	13
Australia	354	3.53%	40	17
Canada	292	2.91%	26	49
France	249	2.48%	6	12
Italy	243	2.42%	24	23
India	209	2.08%	112	180
Spain	181	1.80%	16	27

Notes: The table delineates the distribution of publications on sustainable development topics, categorized by the corresponding author's country of affiliation. The 'Publication Count' column reflects the number of publications attributed to each country, while the 'Percentage of Publications' column represents the proportion of each country's publications relative to the total identified UN SDG-related papers, amounting to 10,031. The 'SDGI Rank' and 'EPI Rank' columns indicate each country's global standing based on official statistics from the Sustainable Development Goals Index and the Environmental Performance Index, respectively.

Table 5: Publication and Citation Counts in Finance Journals

Source title	Publication Count	Citation Count
Emerging Markets Finance and Trade	696	5675
Finance Research Letters	572	7931
Journal of International Money and Finance	469	7977
Journal of Banking and Finance	456	18440
International Journal of Finance and Economics	381	2376
Journal of Corporate Finance	355	12569
International Review of Financial Analysis	348	6295
Research in International Business and Finance	341	4895
Journal of Sustainable Finance and Investment	303	1881
Journal of Financial Economics	283	17740

Table 6: Publication Count By ABS Category

UN SDGs No.	Thematic Terms	ABS 1	ABS 2	ABS 3	ABS 4	ABS 4*
1	Poverty	74	84	44	13	16
2	Hunger	95	51	42	3	1
3	Health	14	41	26	7	6
4	Education	58	67	49	11	7
5	Gender Equality	104	129	142	59	21
6	Water and Sanitation	2	3	3	0	0
7	Energy	121	164	153	7	5
8	Economic Growth	583	1224	1396	254	178
9	Infrastructure	97	198	144	26	21
10	Inequality	144	262	397	111	79
11	Cities	62	169	152	42	37
12	Sustainable Consumption	179	320	365	120	96
13	Climate Change	81	187	123	15	12
14	Oceans	6	2	3	0	0
15	Land Life	1	2	0	0	0
16	Peace and Justice	159	393	506	149	95
17	Partnerships	57	233	287	27	13

Table 7: Percentage Distribution of Thematic Terms Across ABS Categories

UN SDGs No.	Thematic Terms	ABS 1	ABS 2	ABS 3	ABS 4	ABS 4*
1	Poverty	4.03%	2.38%	1.15%	1.54%	2.73%
2	Hunger	5.17%	1.45%	1.10%	0.36%	0.17%
3	Health	0.76%	1.16%	0.68%	0.83%	1.02%
4	Education	3.16%	1.90%	1.28%	1.30%	1.19%
5	Gender Equality	5.66%	3.66%	3.71%	6.99%	3.58%
6	Water and Sanitation	0.11%	0.09%	0.08%	0.00%	0.00%
7	Energy	6.59%	4.65%	3.99%	0.83%	0.85%
8	Economic Growth	31.74%	34.68%	36.43%	30.09%	30.32%
9	Infrastructure	5.28%	5.61%	3.76%	3.08%	3.58%
10	Inequality	7.84%	7.42%	10.36%	13.15%	13.46%
11	Cities	3.38%	4.79%	3.97%	4.98%	6.30%
12	Sustainable Consumption	9.74%	9.07%	9.53%	14.22%	16.35%
13	Climate Change	4.41%	5.30%	3.21%	1.78%	2.04%
14	Oceans	0.33%	0.06%	0.08%	0.00%	0.00%
15	Land Life	0.05%	0.06%	0.00%	0.00%	0.00%
16	Peace and Justice	8.66%	11.14%	13.20%	17.65%	16.18%
17	Partnerships	3.10%	6.60%	7.49%	3.20%	2.21%

Table 8: Chi-square test results: SDG topics of paper distribution in ABS classification

Notes: This table presents the chi-square statistics for each thematic term under the UN SDGs across different Academic Journal Quality Guide (ABS) categories. The values indicate the extent of deviation of the observed distribution of papers from the expected distribution under the null hypothesis, which assumes a uniform distribution across ABS categories. A higher chi-square value suggests a greater disparity in the representation of a specific UN SDG theme within the ABS categories. The degrees of freedom = (Number of ABS categories - 1) * (Number of UN SDG themes - 1)= (5 - 1) * (17 - 1) = 4 * 16 = 64. The overall chi-square test, with a total statistic of 638.346315 and a critical value of 93.22 at a 0.01 significance level for 64 degrees of freedom, led to the rejection of the null hypothesis.

UN SDGs No.	Thematic Terms	ABS 1	ABS 2	ABS 3	ABS 4	ABS 4*
1	Poverty	29.09	0.70	18.53	1.56	0.82
2	Hunger	115.16	2.55	10.70	9.84	8.70
3	Health	0.31	3.07	1.84	0.03	0.13
4	Education	18.56	0.17	5.91	1.18	1.22
5	Gender Equality	8.18	3.22	2.96	14.48	0.68
6	Water and Sanitation	0.28	0.04	0.00	0.64	0.44
7	Energy	24.03	1.43	0.53	23.10	15.86
8	Economic Growth	3.26	0.24	5.58	4.16	2.58
9	Infrastructure	2.01	8.32	5.56	4.11	1.27
10	Inequality	4.44	13.90	4.25	13.11	10.64
11	Cities	3.99	1.59	1.27	0.77	5.17
12	Sustainable Consumption	0.31	4.15	1.52	13.67	22.16
13	Climate Change	1.06	16.75	5.09	9.97	5.32
14	Oceans	8.84	0.75	0.24	0.87	0.61
15	Land Life	0.45	1.01	1.08	0.24	0.17
16	Peace and Justice	19.37	3.57	2.85	20.13	7.42
17	Partnerships	23.10	3.87	18.74	9.87	13.03

47

Table 9: Topic Labels by LDA

Topics No.	Topic Label
1	Financial Risk and Valuation
2	Market Dynamics and Volatility
3	Economic Policy and Rates
4	Corporate Finance and Banking
5	Corporate Governance and Performance
6	Green Economic

48

Table 10: Summary of LDA Topics and Keywords

Topic No.	Topic Label	Ten Topic Keywords
1	Financial Risk and Valuation	Risk, Insurance, Financial, Value, Model, Approach, Market, Portfolio, Asset, Price
2	Market Dynamics and Volatility	Market, Financial, Stock, Markets, Crisis, Global, Policy, Exchange, Price, Volatility
3	Economic Policy and Rates	Rate, Model, Consumption, Income, Tax, Financial, Credit, Policy, Economy, Rates
4	Corporate Finance and Banking	Firms, Banks, Firm, Csr, Bank, Corporate, Effect, Higher, Credit, Information
5	Corporate Governance and Performance	Performance, Governance, Esg, Board, Corporate, Social, Firms, Funds, Gender, Firm
6	Green Economy	Financial, Growth, Economic, Development, Countries, Green, Purpose, Approach, Energy, Real

49

Appendix C - Explanation of Fuzzy Match and its Application in This Research

(1) Fuzzy Match Principle

The principle of Fuzzy Match rests on the concept of approximate string matching, which is a method used in computer science to identify strings that match a pattern approximately rather than exactly. In the domain of text analysis, this is essential because textual data is often riddled with noise—spelling variations, synonyms, abbreviations, and other inconsistencies. Fuzzy matching techniques use algorithms to determine the similarity between strings, enabling the system to find matches that are not 100% identical but are close in context or meaning.

One common approach to fuzzy matching is the use of edit distance metrics, such as the Levenshtein distance, which quantifies how dissimilar two strings are by counting the minimum number of operations required to transform one string into the other. Operations can include insertions, deletions, or substitutions of characters. Another approach is token-based matching, where the text is broken into tokens (usually words), and matches are found based on shared tokens and token proximity, allowing for flexibility in word order and presence.

Algorithms also employ matching techniques like phonetic matching, which identifies words that sound similar when spoken; this is particularly useful for matching names and terms in financial texts where different spellings may represent the same phonetic sounds. Moreover, fuzzy logic can be incorporated into these algorithms, which allows for a range of truth values between "completely true" and "completely false," reflecting the degree of match between strings. It enables the algorithm to handle uncertainty and partial truths that are characteristic of natural language.

(2) Application in This Research

In the context of this research, Fuzzy Match is employed to categorize abstracts from financial journals according to their relevance to the UN SDGs. Given the textual variability and the complex, multi-faceted nature of the SDGs, a traditional keyword match could be overly restrictive and fail to capture the essence of the abstracts. Therefore, the Fuzzy Match approach enables the identification of abstracts that align with the SDGs in a more holistic and nuanced manner, even if they do not contain exact keyword matches.

For instance, an abstract discussing "equitable education funding" may not use the exact phrase "Quality Education," which is the title of SDG 4. However, a fuzzy matching algorithm can recognize the relevance based on the context provided by words like 'equitable' and 'education,' even if the exact SDG terminology is absent.

In applying Fuzzy Match, we first faced the challenge of training the model with the right inputs. The original SDG texts, being descriptive, had the potential to diffuse the focus of the match due to an abundance of non-critical words. By using TF-IDF and Rake algorithms, we refined our inputs to include key terms that capture the essence of each SDG. For example, TF-IDF helped us filter out common words and retain terms with high specificity to the SDGs, while Rake identified phrases that are central to the SDGs' goals.

To enhance the matching process, the study implemented a pre-processing step by extracting 10 key feature words from the UN SDGs' texts. This was achieved using TF-IDF and Rake algorithms to distill words and phrases that are most representative of each SDG's thematic content. TF-IDF highlights words that are uniquely frequent in a particular document relative to a corpus, while Rake focuses on identifying

50

key phrases based on their occurrence and co-occurrence with other words.

Furthermore, the integration of the GPT-4.0 model offered a novel feature extraction method. By instructing GPT-4.0 to extract "10 words that best represent the essence of the text", the study leveraged advanced natural language processing to enhance the Fuzzy Match process. This innovative step allowed for the capturing of conceptual themes that traditional methods might overlook. After extracting ten keywords, perform manual screening to delete some obvious misjudgments, and then use them as input values of the model to train the classifier model.

Subsequent testing on a sample of 100 abstracts with these classifier model provided insights into their effectiveness. The accuracy of the Fuzzy Match was gauged by comparing the model's categorization outcomes with expert assessments. This comparative analysis was crucial in determining the most effective strategy for aligning journal abstracts with the relevant UN SDGs. The picture shows the entire process. Based on accuracy matching, this study applied classifier 8 as the classifier for all articles.

In summary, Fuzzy Match served as a sophisticated methodological framework within this research, accommodating the complexity of the UN SDGs and the varied nature of the textual data from financial journal abstracts. Through the compare of traditional algorithms and advanced models like GPT-4.0, the study achieved a more accurate and reliable categorization of research abstracts, contributing to a deeper understanding of the intersection between finance and sustainable development.

[PLEASE INSERT Figure 18 HERE]

51

Appendix D - Explanation of Latent Dirichlet Allocation (LDA) and its Application in This Research

(1) LDA Principle Latent Dirichlet Allocation (LDA) is a generative statistical model that explains a set of observations through unobserved groups that explain why some parts of the data are similar. It is particularly used in natural language processing to identify topics within documents. LDA assumes that documents are a mixture of topics and that each word in the document is attributable to one of the document's topics.

The model operates under the premise that each document can be represented as a probabilistic distribution of topics, and each topic can be represented as a probabilistic distribution of words. LDA uses the Dirichlet distribution to model these distributions over topics and words, which helps in handling the variability and uncertainty inherent in language.

A crucial aspect of LDA implementation is determining the optimal number of topics (K). This selection is typically guided by model evaluation metrics such as perplexity, which assesses how well a probability model predicts a sample; a lower perplexity score implies better model performance. Additionally, the study incorporates heatmap analysis as a complementary method. Heatmaps in LDA context visually represent the relationships between topics, often using color gradients to indicate the degree of similarity or divergence between pairs of topics. This visual representation is instrumental in interpreting the model's output, offering an intuitive understanding of how topics are related and how distinct or overlapping they might be. Such heatmaps are particularly valuable in identifying patterns or clusters within the topic distributions, providing a more nuanced understanding of the thematic structure uncovered by LDA.

(2) Application in This Research In this study, LDA was applied specifically to a dataset of financial literature that had been pre-selected through a prior classifier for its relevance to the UN SDGs. This dataset, composed exclusively of papers pertaining to the intersection of finance and the SDGs, served as the foundation for our LDA analysis. Through this approach, we were able to systematically identify and examine the key themes and topics prevalent in scholarly discussions about finance within the context of sustainable development. This targeted LDA application on a dataset of UN SDG-related financial papers provides valuable insights into the thematic focus and evolving trends in the nexus of finance and sustainability.

To determine the optimal number of topics, we examined the topic perplexity across different models. Perplexity serves as an indicator of the model's predictive power, with the goal being to choose a number of topics that provide the most coherent and distinctive groups of words without overfitting. The analysis indicated a plateau at six topics, which suggested that increasing the number of topics further would result in only marginal improvements in model fit. Thus, six topics were deemed to provide a balance between specificity and breadth.

After identifying six LDA topics, LDA modeling was performed. The model will classify each main document and output 10 keywords. The potential topics of these ten words are then manually found. The following are the keywords of the 6 topics output by LDA, and the topics summarized manually.

[PLEASE INSERT Table 10 HERE]

52

These topics were then matched with the 17 UN SDGs to explore the intersection of finance research themes and sustainable development. The LDA model's ability to distill large volumes of text into distinct themes allowed for a systematic categorization of the research abstracts, facilitating a nuanced analysis of how financial research contributes to, or aligns with, the broader goals of sustainability.

Through this application of LDA, the research provides insights into how topics prevalent in financial journals resonate with the UN SDGs, highlighting not only the thematic concentration within the financial literature but also the sector's engagement with global sustainability issues. This alignment is critical for understanding the impact of finance on society and the environment, and it underscores the importance of integrating sustainable development considerations into financial research and practice.