**Manuscript Number:** ECONJOURNAL-D-24-00080

**Full Title:** Asset Pricing and Portfolio Investment Management Using Machine Learning: Research Trend Analysis Using Scientometrics

**Article Type:** Research Article

**Keywords:** Asset Pricing; Investment Portfolio; Key Indicators; Scientometrics; Machine Learning

**Manuscript Region of Origin:** CHINA

**Abstract:**

"Asset pricing" in the context of financial economics pertains to the investigation and formulation of two fundamental pricing ideas and the models that go along with them. Various models exist for different scenarios, but they can be traced back to either general equilibrium asset pricing or rational asset pricing. Asset pricing models, as the name suggests, serve as valuable tools to assess the value of assets. The general equilibrium theory states that supply and demand interact to determine market prices. In this context, asset prices collectively satisfy the market clearing condition, which dictates that the supply and demand for each asset are equal at the prevailing price. Another crucial aspect of financial planning is portfolio management (PM), which aims to maximize investment profits while minimizing losses. Portfolio management involves implementing effective asset allocation strategies to enhance returns and mitigate risks. Numerous studies have been conducted worldwide on various types of asset pricing models and investment portfolios, with some incorporating machine learning and deep learning techniques. In several models, the predictive accuracy has exceeded 90%. To shed light on the current research landscape in the realm of asset pricing and portfolio investment, we conducted a scientometric analysis.

**Manuscript Classifications:** 7: Financial Economics; 15: Economic Development, Innovation, Technological Change, and Growth

Ting Li 1*

1 Faculty of Intelligent Finance, Henan Institute of Economics and Trade, 450046, Zhengzhou, China.

*Corresponding author email: liting2302@outlook.com (Li Ting)

Abstract: "Asset pricing" in the context of financial economics pertains to the investigation and formulation of two fundamental pricing ideas and the models that go along with them. Various models exist for different scenarios, but they can be traced back to either general equilibrium asset pricing or rational asset pricing. Asset pricing models, as the name suggests, serve as valuable tools to assess the value of assets. The general equilibrium theory states that supply and demand interact to determine market prices. In this context, asset prices collectively satisfy the market clearing condition, which dictates that the supply and demand for each asset are equal at the prevailing price. Another crucial aspect of financial planning is portfolio management (PM), which aims to maximize investment profits while minimizing losses. Portfolio management involves implementing effective asset allocation strategies to enhance returns and mitigate risks. Numerous studies have been conducted worldwide on various types of asset pricing models and investment portfolios, with some incorporating machine learning and deep learning techniques. In several models, the predictive accuracy has exceeded 90%. To shed light on the current research landscape in the realm of asset pricing and portfolio investment, we conducted a scientometric analysis.

Keywords: Asset Pricing, Investment Portfolio, Key Indicators, Scientometrics, Machine Learning
1. Introduction

Financial management is the integration of management concepts to the administration of an organization's finances. As part of finance management, an individual's or organization's finances are strategically planned and managed to better align their financial status with their goals and objectives. Financial management is the process of planning, forecasting, controlling, and managing a company's financial resources to achieve its objective. It aims to maximize the firm's financial management in order to boost investor profits. It addresses every aspect of profitability, expenses, finances, and credit. The cornerstone of any company or individual is prudent money management. They need to outperform the related costs in terms of rate of return on capital sourced from the market. A company's accounting and legal departments are usually balanced by financial management. It is composed of three essential components: bringing down the cost of borrowing, making sure there is adequate money, and making prudent use of that money. In financial management asset pricing and investment portfolio management is studied by many researchers and also organisations.

2. Asset pricing

The formal study and development of the two primary pricing concepts and the ensuing models is what is meant by "asset pricing" in the field of financial economics (Gu et al., 2018; 2020). Asset pricing refers to the method by which the market value of a financial asset is determined, taking into account the asset's level of risk and the return investors can expect to receive on their investment. Since the values of stocks, bonds, derivatives, and other financial instruments are always changing, it is important for investors to have a firm grasp of the principles underlying asset pricing. The Capital Asset Pricing Model (CAPM) and the Arbitrage Pricing Theory are just two of many possible frameworks to use when determining an asset's value (APT). Each model has its own set of pros and cons, but they all share a common goal of helping investors make better choices about where to put their money. Asset pricing is a crucial part of finance, despite its complexity. By understanding how asset prices are determined, investors will be better equipped to make informed decisions regarding which assets to buy and sell. Insights like these can also help people avoid overspending on items whose true worth might not be as high as they were led to believe. General equilibrium asset pricing and rational asset pricing are the two main theoretical foundations from which a wide variety of scenario-specific models have emerged. As their name implies, models are useful tools for determining the value of various assets. In other words, they help us determine how much something is worth. The general equilibrium theory states that the market sets prices in response to supply and demand. In this case, asset prices as a whole satisfy the so-called market clearing condition, which states that the supply and demand for each asset is equal at the current price.
3. Portfolio management

Also, a key component of financial planning is portfolio management (PM), which tries to accomplish investment objectives like maximum returns or minimal risks. By effectively allocating assets, portfolio management is a set of procedures that increases returns and reduces risk. Portfolio management is the process of regularly reallocating funds into financial assets with the goal of increasing the expected return on investment while lowering risk. Additionally, research frequently ignores portfolio construction in favour of prediction accuracy. Furthermore, index data, which hardly ever accurately reflects prices to buy or sell an item, has typically been used to evaluate the methodologies. Exchange Traded Funds (ETFs) are used as the evaluation's base assets to get around these issues, and we provide a two-stage deep learning architecture called Grouped-ETFs Model (GEM) with a joint cost function (Yun, et al., 2020). The investment policy, which entails reallocating the various assets in order to raise the anticipated return on investment (Ye et al. 2020). A delay between the purchase of an asset and its payment could result in settlement risk if the contract's requirements aren't met (Soleymani, 2020).

Investment portfolio management entails analyzing various investment opportunities, selecting and assembling the optimal investment combination to reach the objective, routinely assessing and reassessing the investment portfolio, and implementing the required adjustments. Different investing options carry varying degrees of risk and produce distinct rewards over various time frames. Consequently, a portfolio manager decides to invest in a variety of options in order to reduce risk and maximise earnings. The term "investment portfolio" refers to this collection of carefully chosen investments. Figure 1 indicates the different types of the investment portfolio process.

**Aggressive portfolio:** made up of high-risk investments, including commodities, futures, options, and other securities, with the expectation of quick profits.

**Defensive Portfolio:** A defensive portfolio is made up of low-risk stocks with consistent earnings, such as investments in blue-chip stocks of excellent quality.

**Hybrid portfolio:** The most well-balanced and frequently utilised portfolio by the portfolio manager is the hybrid one. It is possible to see a good mix of various asset types, some of which have high risk-high return profiles and others that have low risk-poor return profiles.
Investment portfolio management variables
Planning an investment portfolio requires taking into account a number of variables relating to the investor's characteristics. The following is a discussion of the key variables affecting investment portfolio management:

**Time Frame:** The length of time the investor is willing to put money into an investment determines the type of investment made. Stock and equity investments must be made long-term to provide large returns.

**Age of Investor:** An investor's age affects the type of investment, their capacity for risk, and the yield on their investments. A youthful investor can be able to pool money for long-term investments and take a high risk at the same time to generate better profits, and vice versa.

**Risk Tolerance Level:** An investor's investment portfolio is influenced by the amount of risk they are willing to take. Those with modest incomes or elderly investors may not be prepared to invest in high-risk assets. Research related to asset pricing and portfolio management for investment is steadily growing and is considered as one of the leading research avenues for management researchers.
4. Scientometrics-Based Research Trend Analysis

The study of the numerical aspects of the scientific communication process is known as scientometrics. In this research, we examine sources of citation data, journal impact factors, and the "rules" of scientometrics. The authors advocate for integrative and interdisciplinary research techniques, for strengthening fundamental, methodological, and experimental research programmes, for independent funding of research, and for an improvement of the databases.

4.1 Evolution of scientometrics as a field

Scientometrics is a term coined by Vassily V. Nalimov to describe the growth, organisation, interrelationships, and productivity of research. Bibliometrics and informetrics share similarities and areas of study with scientometrics (Nalimov, 1971). Study topics include tracking variations in scholarly production through time and within nations. It is widely used in many domains and verticals including engineering, finance, management, arts, culture, humanities and other domains (Pena et al., 2023; Gillet et al., 2023; Nyika et al., 2022; Umeokafor et al., 2022). Scientometrics is applied to wide areas in finance, marketing, management and other areas, recently there has been an impetus in this field with many researchers working on using it for analysing research trends (Baker et al, 2021; Cai & Guo, 2021; Costa et al., 2017; Fat et al., 2010; Maia et al., 2019).

4.2 The evolution of scientometrics across time

The study of science and technology's quantitative aspects as a discourse form is known as scientometrics (Garfield, 1955). The key themes cover how to assess the value and impact of research, how to comprehend the citation process, and how to map scientific areas. In addition to other closely related subjects, scientometrics is concerned with communication in the social sciences, arts, and sciences (Hood et al., 2001; Mingers et al, 2015). The topic is primarily concerned with the analysis of web pages as documents. Scholarly impact metrics based on activity in online tools and environments are studied by a field called altmetrics. This section replaces journal citations with the impact of social networking tools, such as views, downloads, and "likes". Scientometrics is the area of study and evaluation of scientific research that is most closely related to it asset pricing and investment portfolio model development (Glänzel et al., 1994; Yu et al., 2015; Wouters, 2014). Webometrics and altmetrics have advanced as due to the breakthroughs in this subject; these topics will be examined later. The development of scientometrics has revolved mostly around one central idea: citation (Garfield, 1979). The citation establishes the associations between people, ideas, publications, and organizations that are necessary to build a network or empirical field that can be subjected to quantitative analysis. It also offers a chronological link between the earlier publishing of the references and the later appearance of the citations.
4.3 Indicators of scientific productivity

Some of the earliest studies, dating back to the 1920s, examined productivity in terms of the quantity of papers authored or produced by a research unit (Larivière, 2006). The paretian phenomenon holds that a small percentage of producers account for a large percentage of outputs (Priem et al. 2014). The original works should be acknowledged as being illustrative and offering few instances (Lotka, 1926; Price, 1963). The majority of scientometrics study uses either Web of Science or Scopus database to reproduce the productivity of the particular topic (Amara & Landry, 2012). Majorly Scopus is widely used for this study due to its popularity and preciseness. In this work, we have used the Scopus database for the study purpose, which index many journals covering the keyword.

5. Methodology

The methodology is broken down into three distinct parts, including an initial phase in which we analyse the entire body of research being conducted with multiple keywords before narrowing in on "Asset Pricing," "Investment Portfolio," and "Machine Learning" as the most appropriate. In this analysis, we only included years with actual findings, ignoring years with no findings. The next step was to use the following specified approach to collect data from Scopus. The final step is to analyse the data gathered using tools like Biblioshiny and VoS Viewer. The gathered data was then analysed using a scientometric method, and conclusions were drawn.

Using scientometric analysis, relevant literature was uncovered and assessed for this work. Scopus, one of the largest citation databases, is used in this research. Scopus is managed by Elsevier, Inc. Scopus, Elsevier's database of abstracts and citations was launched in 2004. It has good coverage of articles, including 37000 titles from 11500 publishers, and almost all are peer-reviewed, from prestigious publishers and academic societies. Scopus indexes a large number of reputable journals across numerous disciplines, including the health sciences, social sciences, and physical sciences, as well as to life sciences. The scope of this term includes periodicals, book series, and journals. The quality of each journal in the Scopus database is checked annually using a variety of quantitative metrics such as the h-Index, CiteScore, SJR (SCImago Journal Rank), and SNIP (Source Normalized Impact per Paper). Scopus queries can also look through patent databases. Trends in the publication of academic works like research papers, conference proceedings, and other scholarly documents are analysed quantitatively and statistically using a tool called scientometrics. Extensive searches of Scopus databases were performed on October 3, 2022. (www.scopus.com). In order to mitigate the risk for bias introduced by daily modifications to the databases, the search was carried out on a singular day. Scholarly works indexed by Scopus and published between 2013 and 2022 were chosen for this retrospective study. More so than other databases like Web of Science, Dimensions and etc. Scopus is often used in scientometric and bibliometric research because of the sheer volume of data it contains. Scopus has gained acclaim as a trustworthy and extensive database of scholarly
articles. It has become increasingly popular as a bibliometric data source, with numerous studies citing it. The objective of this study is to present a comprehensive scientific synopsis of the numerous institutions, organizations, and researchers that have collaborated to improve our knowledge of portfolio management and asset pricing via porous media. The acquired data was also analysed using tools like keyword analysis, co-occurrence networks, conceptual structure maps, three-factor analysis of countries, authors, and keywords, and thematic development.

5.1 Search Query

To obtain scientometric information on the selected topic we first generated keywords related to this domain, done keyword research and finalised the following keywords, it was challenging to obtain the keywords of an upcoming domain, and it was done by carefully verifying multiple articles.

5.2 Search criteria

(TITLE-ABS-KEY (asset AND pricing) OR TITLE-ABS-KEY (portfolio AND investment) AND TITLE-ABS-KEY (machine AND learning)) AND (EXCLUDE (PUBYEAR, 1998) OR EXCLUDE (PUBYEAR, 1997) OR EXCLUDE (PUBYEAR, 1994) OR EXCLUDE (PUBYEAR, 1971))

5.3 Criteria of Inclusion and Exclusion

After the first search using the keywords, we are able to obtain 510 papers using Scopus. When the time frame for publication was constrained to be between the years 2000 and 2023, a total of 481 works were found. The findings were narrowed to 471 books, book chapters, conference papers, reviews, and articles that addressed asset pricing and portfolio management in porous media and have been published in the top 10 countries for this subject. Only materials written in English will be evaluated. From an initial set of 3066 documents, 2026 were removed after strict criteria for all admissible documents were applied. From 2013 to 2022, a total of 1040 records, including articles (N=794), conference papers (215), book chapters (21), review articles (7), and books (3), were published. This paper provides a hybrid strategy, combining scientometric and complex network analytics, to thoroughly examine the literature on asset pricing and portfolio management via porous media. Figure 2 depicts a research workflow and the integration of analytical tools. Two of the writers in this paper followed the same procedures to double-check the items chosen and the accuracy of the data.
5.4 Data Analysis

Several bibliometric analysis techniques were used to sift through the 481 generated documents and pull out the most relevant data. Microsoft Excel (Version 2109) was used to perform a number of fundamental duties, including research into publishing and citation patterns. The authors employed numerous software programs to facilitate the visualisation process, such as Biblioshiny (version 4.1.1), OriginPro 2022b (64-bit), and VOS viewer (version 1.6.15). These tools also helped reveal relationships in co-authorship, co-citation, and bibliographic coupling.

6. Results & Discussion

6.1 Analysis of the Overall Growth Trend

Below shown figure provides comprehensive details about the research progress in the selected keywords through scientometric analysis. The terms mentioned below indicate parameters which are time span of research, sources in total, documents obtained for the keywords, the annual growth rate in %, number of authors working towards this area, how many single-authored papers (this shows how strong this domain is), internationally co-authored papers in %, how many co-authors are there per document, how many keywords authors mentioned in total, how many articles were references, documents averaged and average citations. These are the important parameters that describe the growth of research in this particular area. Figure 3 shows the summary of the output obtained via this study.
Figure 3: The main results of the scientometric analysis

The above-mentioned figure demonstrates the publications frequency and also citations based on keywords, asset pricing and portfolio management obtained through Scopus for the period of 2000-2023. It is very evident that the publications and successive citations were increased steadily over the years from the starting period. Also, it is clear that the most successful year is recent year (2022) in which more than 110 articles were published and also citations reached a peak value of 1510 (taken on October 2022). In 2019 there is a small drop in number of publications which is attributed to the Covid onset which put pressure on complete academic industry. This trend of drop is steadily reversed during next few years which shown steady raise in the number of publication and citations. It is also found that comparatively this domain gets less citation when compared with other critical areas of research in management and finance and this point need to be considered while undertaking newer research. Figure 4 represents the publication and citation trends occurred during the study period.

Figure 4: Publication and citations trend (2000-2022).
6.2 Top Ten Publishing Countries

In Table 1 we see the top 10 most productive nations. When compared to the United States (184 publications), which came in second, and the United Kingdom (36 publications), which came in third, it is clear that more research needs to be done in the areas mentioned in this article. The most productive country was China, with 114 publications. The United States holds the lead in citations with a total of 1356, indicating that its researchers are more extensively cited and read compared to their counterparts in other nations. China comes in second with 815 citations, which is also very respectable. The United Kingdom also contributed significantly, with 526 citations; the fact that citations from other countries are still rising demonstrates the field's status as a promising new area with ample room for study.

<table>
<thead>
<tr>
<th>Country</th>
<th>Documents</th>
<th>Citations</th>
<th>Percentage Contribution</th>
<th>Citation Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>114</td>
<td>812</td>
<td>20.81</td>
<td>7.12</td>
</tr>
<tr>
<td>United States</td>
<td>112</td>
<td>1356</td>
<td>34.75</td>
<td>12.11</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>36</td>
<td>526</td>
<td>13.48</td>
<td>14.61</td>
</tr>
<tr>
<td>India</td>
<td>35</td>
<td>178</td>
<td>4.56</td>
<td>5.09</td>
</tr>
<tr>
<td>Germany</td>
<td>27</td>
<td>103</td>
<td>2.64</td>
<td>3.81</td>
</tr>
<tr>
<td>Italy</td>
<td>23</td>
<td>240</td>
<td>6.15</td>
<td>10.43</td>
</tr>
<tr>
<td>Taiwan</td>
<td>22</td>
<td>291</td>
<td>7.46</td>
<td>13.23</td>
</tr>
<tr>
<td>France</td>
<td>14</td>
<td>223</td>
<td>5.72</td>
<td>15.93</td>
</tr>
<tr>
<td>Brazil</td>
<td>12</td>
<td>110</td>
<td>2.82</td>
<td>9.17</td>
</tr>
<tr>
<td>Canada</td>
<td>12</td>
<td>63</td>
<td>1.61</td>
<td>5.25</td>
</tr>
</tbody>
</table>

Figure 5 highlights the top 10 publishing countries in the field of asset pricing and portfolio management through porous materials and their corresponding citations during the period 2013-2022. In any research, it is mandatory to understand the performance of each country involved in the research and this study provides insight into the same.
Figure 5: The top ten publishing nations and their corresponding citations
6.3 Top Ten High-Research-Producing Institutions

Figure 6 lists the top 10 research-producing companies. More than five articles were produced by each of the top ten organizations; this indicates that the research in this particular field continues to be scattered rather than concentrated at one or two domains. Stevens Institute of Technology and Xi’an Jiaotong-Liverpool University had produced maximum papers (9) in these particular keywords and showing considerable trend in performing more research. This is followed by University of Chicago with 6 publications to its count which other mentioned institutes follow in the graph. With this it is very evident that more research should be undertaken in the machine learning application for asset pricing and investment portfolio management.

![Figure 6: Top ten most highly productive organizations.](image)

6.4 Most Prolific Authors

According to the data, from January 2013 through June 2022, the ten most productive researchers in the field of asset pricing and portfolio management are shown in figure below. These scientists routinely advance our understanding of this topic. Results showed that Zhang Y and Lu J, Chinese researchers, produced more documents and was cited most frequently. Huang CF has a high h-index which indicates that he is, to this day, the most cited researcher in the study of how waves travel through porous media. Concerned with this particular work there are not much papers available and the highest paper by an author seems to be only 4 which is well below the average level. Hence, it is very clear that more research in this area is necessary considering the need and demand available in the field, we recommend more authors to involve in active research in this field. Liu J and Li J of China is the author who consistently performed in this research and continuing publications with good number of citations. Many authors had published 3 papers in this domain totally for the period of this research which can be taken as
range of publication. Therefore, it is confirmed that they have made a major contribution to our understanding of how to perform machine learning based research focussing on asset pricing and portfolio management modelling. Additional seminal researchers in this field are shown in Figure 7 and Figure 8.

The line in below mentioned figure represents the author's writing process. The size of the bubbles indicates the number of documents that author produces annually, and the color intensity of the bubbles indicates the total number of citations. When an author first began contributing to a particular field, their publication history is represented by the first bubble on the line. The greater the size of the bubble, the more articles that writer produces annually. The greater the depth of the bubble's colour, the greater an author's annual citation rate.

![Top-Authors' Production over Time](image)

**Figure 7:** Top ten authors’ production over time
Figure 8: Author impact by H-Index

6.5 Three-Factor Analysis

6.5.1 Keywords, Authors and Countries

Figure 9 provides a better visual depiction of the nations and areas of interest for the active researchers in the field of porous media asset pricing and portfolio management. The figure is a three-field plot of country, author, and keyword contributions to the literature on the topic of machine learning applications in asset pricing and portfolio management studied during the mentioned period of time.

The first column features a list of countries with recent activity, the second features a list of scholars from those countries, and the third features a list of keywords that have been used frequently by the authors. Pay attention to the distance between the boxes and the width of the lines within them. The list shows numerous authors have some connection to China, making it the country with the most author affiliations overall. After that comes USA, United Kingdom and Hong Kong. The thick line connecting countries and authors indicates that Li J. and Zhang Y. continue to be the most significant contributors from China. The two most prolific writers in the USA are Yang H and Zhang I. The same applies to China, where Li and Zhang are essentially the authors who have made significant contributions to the study of asset pricing and portfolio management. The curved line in this diagram represents the writer's actual handwriting. The larger the bubbles, the more documents that author produced that year; the darker the colours, the
more times that document was cited. The author's major contributions to the field under consideration are indicated by the first bubble on the line. The writer produces more articles annually the larger the bubble. The darker the bubble, the higher the annual citation rate of the corresponding author.

Figure 9: A three-field plot of countries, authors, and topics of asset pricing and portfolio investment

6.5.2 Thematic Evolution based on Keywords

The evolution of keywords in three distinct stages (2000–2010, 2011–2020, and 2021–2023) is shown in Figure 10. It is apparent that the majority of the second stage's research focuses on porous materials, asset pricing, and portfolio management. The height of the boxes and the thickness of the connecting lines between 2021 and 2023, which show that there is a comparatively greater focus on these two terms in the research. With this information, one can easily identify what sort of keyword is being used in this particular study and can persuade research in that direction.
6.5.3 Thematic Map

In addition, this study thematically analysed asset pricing and portfolio management through porous materials. A thematic map's objective is to assess a field's long-term viability as well as its current state. The results of this analysis can be used to inform researchers and other interested parties about the prospects for developing subfields within a field of study. Thematic analysis is a method for discovering overarching themes by examining related groups of keywords. These subjects exhibit some similarities (density and centrality). Centrality is shown along the horizontal axis, and density is depicted along the vertical axis. Density evaluates the network's cohesiveness, while centrality assesses the strength of connections between ideas (Esfahani et al., 2019). Both factors are used to determine how significant and advanced a subject is. In a theme network, the more connections a node has to other nodes, the more important it is. This is comparable to how a research field's capacity for growth and sustainability is determined by the cohesiveness of a node, which stands in for the density of the field. The region of asset pricing and portfolio management through porous materials (Q1 to Q4) is shown in Figure 11, a four-quadrant thematic map.
Figure 11: Thematic map: The primary theme is in Q1, while more in-depth and specialised themes in Q2 connect to the primary theme; topics that are disappearing or emerging in Q3; Foundational and cross-cutting themes make up Q4.

Q1 in the top right represents the primary themes, Q4 in the lower right represents the secondary themes, Q2 in the upper left represents the very specific themes, and Q3 in the lower left indicates the themes that are either emerging or fading away. The concept of "data handling and economic analysis," which is wedged between Q1 and Q4, is sufficiently developed to provide a framework for the study of this area. With this keyword and themes, asset pricing and portfolio management research can be deployed, and it will always be the primary focus. Themes such as "algorithm", "intelligence", "performance assessment," "price dynamics" and "product" seen in Q1 are the basics and are critical for the field’s development. Q2 themes contain internal links but contribute little to this particular field. According to this research, themes from Q2 that could benefit from further connections to porous materials include the financial market, genetic algorithms, stock selection, soft computing, financial data processing, and portfolio management. In order to offer asset pricing solutions appropriate for the twenty-first century, scholars in this discipline might investigate several themes.

The theme of Q3, which demonstrates coherence with the themes of Q4, suggests that some of its elements are fundamental and essential for the advancement of the study on this domain. The themes of Q3, risk perception, financial market, commerce, machine learning, and costs are in use but still in the early stages of development. According to the thematic analysis, further work is needed to create themes like "portfolio assessment" and its related concepts like "product portfolio" and "competition," in order to forge stronger connections with "managers."
6.5.4 Thematic focus of the field of asset pricing and portfolio management through porous materials

(a) Keywords analysis, co-occurrence network, and Conceptual Structure Map

The topics that have received the most attention from researchers studying asset pricing and portfolio management through porous materials between 2013 and 2022 are discussed in this section. The study also seeks to clarify whether or not the topic of debate among subject-matter experts has shifted. Initially, the frequency with which a set of keywords appeared is analysed. The dynamics of keywords, popular subjects, the co-occurrence network, and thematic subfields were then examined.

The investigation of emerging themes and the research interests of experts in an area is facilitated greatly by the examination of the keywords that authors of published works use (Song et al., 2019). This research is conducted in this way because publication keywords assist in rapidly determining the subject matter and the primary focus of a certain publication. In the papers relating to asset pricing and portfolio management through porous materials, the word-cloud shown in Figure 12 displays the most frequently used keywords.

Figure 12: A visualized Word-cloud of frequently used keywords in the asset pricing and portfolio investment research: These are some of the most frequently used keywords in the field.

Figure 13 displays plus keyword word dynamics. Asset pricing and portfolio management rose rapidly from 2013 to 2022. Scholars discussed asset pricing and portfolio management through
porous material the most in these years. However, commonly used keywords such as financial data processing, deep learning, learning systems, portfolio optimization, machine learning, electronic trading, decision-making, and costs continue to grow in popularity (see Figure 13). This discovery implies that the field of asset pricing and portfolio management will further investigate these prevalent aspects.

Figure 13: Visualized word dynamics of keywords

VOSviewer's occurrences attribute details on how many documents contain a given keyword's occurrences. Co-occurrence networks, a type of graph, are used to display the number of times two keywords appear together. Constructing a co-occurrence network involves assigning a node or point to each variable. Connected nodes share a common keyword. Keywords network paints a clear picture of a research field, making it simple to grasp the topics discussed and the relationships between them. Co-occurrence-minimum keywords were included. Twenty or more of the chosen keywords are adequate. The size of a node correlates with how often it appears in the graph. When a line links two or more nodes, it indicates that they appeared in the same source. When two nodes are closer together, that means the terms are frequently used together.

Table 2: Top keywords with maximum occurrences
Table 2 shows that the top five keywords are investments, financial markets, commerce, electronic trading, and forecasting totalling more than 600 occurrences. This data is obtained from the author keywords used in Scopus and not indexed keywords that Scopus uses to index the article. While performing new research, these keywords can be taken as directional coefficients to perform the search.

This is the co-occurrence network seen in Figure 14. Three clusters, each with a distinct color, were created using the most popular keywords. Clusters were created by grouping keywords with comparable content together. For example, the keywords "financial markets", "costs", "machine learning", "forecasting" and "commerce" were in the red cluster and the keywords "electronic trading", "learning algorithms" and "deep learning" were in the blue cluster, whereas “Porous medium” is in red cluster, “investments” “portfolio managements”, “financial data processing”, “learning systems” and “artificial intelligence” were in green cluster. The thickness of the lines indicated the strength of the co-occurrence of keywords both within and between clusters, while the size of the circles indicated the density of keywords. The image illustrates how strongly the three clusters were related to one another and how each cluster was interconnected. The networked keywords used in this study demonstrated the strong interconnection of several aspects of asset price and portfolio management.

In this network, there were over 300 linkages, 3 clusters, 14 keywords, and a total link strength of 5937. "Investments" is the largest cluster, cluster 2 (green) with five keywords; "financial markets" is the representation of cluster 1 (red) with five keywords; and "learning algorithms" is the representation of cluster 3 (blue) with three keywords. These keywords

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Frequency</th>
<th>Total link strength</th>
<th>Keyword</th>
<th>Frequency</th>
<th>Total link strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>artificial intelligence</td>
<td>57</td>
<td>216</td>
<td>learning systems</td>
<td>112</td>
<td>498</td>
</tr>
<tr>
<td>commerce</td>
<td>141</td>
<td>642</td>
<td>machine learning</td>
<td>260</td>
<td>763</td>
</tr>
<tr>
<td>costs</td>
<td>62</td>
<td>247</td>
<td>machine-learning</td>
<td>53</td>
<td>224</td>
</tr>
<tr>
<td>deep learning</td>
<td>65</td>
<td>294</td>
<td>portfolio managements</td>
<td>58</td>
<td>281</td>
</tr>
<tr>
<td>electronic trading</td>
<td>94</td>
<td>472</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>financial data processing</td>
<td>74</td>
<td>311</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>financial markets</td>
<td>155</td>
<td>724</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>forecasting</td>
<td>86</td>
<td>394</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>investments</td>
<td>264</td>
<td>1019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>learning algorithms</td>
<td>64</td>
<td>301</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Frequency</th>
<th>Total link strength</th>
<th>Keyword</th>
<th>Frequency</th>
<th>Total link strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>artificial intelligence</td>
<td>57</td>
<td>216</td>
<td>learning systems</td>
<td>112</td>
<td>498</td>
</tr>
<tr>
<td>commerce</td>
<td>141</td>
<td>642</td>
<td>machine learning</td>
<td>260</td>
<td>763</td>
</tr>
<tr>
<td>costs</td>
<td>62</td>
<td>247</td>
<td>machine-learning</td>
<td>53</td>
<td>224</td>
</tr>
<tr>
<td>deep learning</td>
<td>65</td>
<td>294</td>
<td>portfolio managements</td>
<td>58</td>
<td>281</td>
</tr>
<tr>
<td>electronic trading</td>
<td>94</td>
<td>472</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>financial data processing</td>
<td>74</td>
<td>311</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>financial markets</td>
<td>155</td>
<td>724</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>forecasting</td>
<td>86</td>
<td>394</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>investments</td>
<td>264</td>
<td>1019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>learning algorithms</td>
<td>64</td>
<td>301</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
indicate the primary research being undertaken in this particular area of research, their co-occurrence indicates the common research progress in this domain.

Figure 14: Keyword co-occurrences with a minimum of 50 co-occurrences; of the 3110 keywords 14 meet the threshold.

By regionally mapping the relationships between phrases, we were able to create a conceptual structure map that depicted the contextual structure of commonly used words in studies of the framing effect (Figure 15). Each word's placement is determined by its Dim 1 and Dim 2 values. In the field of bibliometrics, a "diminutive particle," or "dim" for short, is what creates the connections between terms with similar weight. We used multiple correspondence analysis (MCA), a robust technique for dealing with large datasets. High-dimensional categorical data and K-means clustering can be used to find unrelated collections of texts that share common themes among them, both of which could be part of the complex data used to establish the conceptual framework of a field. The principal component analysis is a form of multidimensional scaling frequently used by researchers when the variables to be analysed are categorical rather than numeric. Groups of variables that are positively correlated are plotted together in multi-
category analysis (MCA), while groups of variables that are negatively correlated are plotted in opposite quadrants of the plot origin. Both the red and the blue sections of this map comprise terms that are connected to one another. The red area contained a large number and variety of words, indicating that many scholarly works provided connections between the terms mentioned there (the top five being "financial markets," "commerce," "machine learning," "forecasting", and "costs").

![Conceptual Structure Map - method: MCA](image)

**Figure 15:** Conceptual Structure Map. Source: Biblioshiny

(b) *Authors bibliometric coupling*

Figure 16 presents the bibliographic coupling of authors in asset pricing and portfolio management through porous materials research. Authors with at least 4 documents and 10 citations were considered; only 9 authors have met the criterion. Authors with the highest bibliographic coupling activity included Zhang H, Hu Y and Liu J who contributed to the maximum bibliometric coupling. Four distinct clusters are used to group authors. Details of these clusters to analyse the authors coupling. Authors who are working in collaboration with one another always increase the research impact via combined citations and also by increased paper count.
(c) Bibliographic Coupling of Countries

The bibliographic coupling of countries in asset pricing and portfolio management is seen in Figure 17. 24 of the 76 countries that met the requirements had a minimum of 5 documents and 5 citations. China, the United States, the United Kingdom, Germany, and Italy had the highest bibliographic coupling activity. Clusters of countries are formed based on the similarities of their contents, and quantitative network metrics characterise the links between clusters. There were four distinct groups identified in the bibliographic linkages of countries studying asset pricing and portfolio management through porous materials. China, is the country that performs the most work in the field of asset pricing and portfolio management through porous materials research, with a good number of publications in asset pricing and investment portfolio research.
Figure 17: Bibliographic coupling of countries that published a minimum of 5 documents and 5 citations; 24 meet the threshold out of 76 countries.

(d) Bibliographic Coupling of Journals

If two journals both reference the same third source, we say that they are bibliographically related. The degree to which a set of publications share a common subject through bibliographic coupling. The bibliographic connection between a few chosen journals is shown in Figure 18. The circle's size and colours represent bibliographic coupling levels and clusters. Journals having at least two articles and twenty citations were included; 14 out of 318 journals met this requirement. Journals with the highest bibliographic coupling activity included Financial data sciences, Review of financial studies, ACM international conference proceedings series, Expert systems with applications, Decision support systems and IEEE access. Clusters are groups of journals with similar content, and quantitative network indicators can be used to describe the relationships between them. By analysing the relationships between journals in the fields of porous materials research and asset pricing and portfolio management, we were able to divide the literature into six distinct groups, which are shown in Figure 18 of the VOSviewer.
Figure 18: Bibliographic coupling of journals with minimum of 2 documents and 20 citations: 14 meet the threshold out of 318 journals

7. Limitations and Future Research Directions

The study only took into account books, reviews, book chapters, conference papers, and articles that were indexed in Scopus. Incorporating other databases (like Web of Science) may help further this work and improve our comprehension of the study of asset pricing and portfolio management through porous materials. Although the researchers made every effort to incorporate every potential search term when shortlisting the articles for the study, more analysis might be conducted by incorporating additional search terms as shown by the keywords cluster presented in Figure 9. Further this article considered only the documents in the final stage of publication and the language as English. The use of a single database (Scopus) and the inclusion of the more general terms "asset pricing AND portfolio management" are limitations of the study. Results from this study are only quantitative and based on correlation, thus qualitative analysis of the data could improve findings. Also, we have used only Scopus as database for this work, similar databases including Web of Science, Dimensions and other related database also exists. It is valid to note that every citation database adopts a unique method to source, segregate and add articles, citations in their service and analysing those databases may provide more clear inputs.
8. Conclusions

The primary goal of this investigation was to review the existing literature on the topic of asset pricing and portfolio management using ML. Research into this area is very much essential for managing individual and organisational assets. For the period from 2000 to 2023, there has not been a single bibliometric analysis that comprehensively evaluates and summarises the literature, progress, and future directions of this major subfield of porous materials.

However, the present article makes a number of recommendations for possible future research topics based on the review and results of the current investigation. Researchers working in the field of asset pricing and portfolio management through porous materials can use this study to develop strategies based on the topics that are emerging (as evidenced by the data visualisation in this study). For the purpose of identifying research gaps and fresh perspectives, they can also identify the most significant publications, authors, and journals in this field.

In conclusion, this research has contributed significantly to the expanding knowledge base on asset pricing and portfolio management through porous materials. This study offers fresh insights into the research field of asset pricing and portfolio management through porous materials by evaluating 481 articles, conference papers, and book chapters from Scopus. Results indicate that from 2013 to 2022, academic interest in this subject gradually grew. The current study, which is the first scientometric study in asset pricing and portfolio management through porous materials, will aid researchers in identifying prior research trends, research gaps, and research implications.

Statements of Declaration

Funding: There is no funding available for this study.

Conflict of interest: The author has no relevant financial or non-financial interests to disclose.

Ethics approval: Not applicable
References


