Implications of AI usage for financial stability: Evidence from AI-driven investment funds identified by generative AI

Victor Pak-Ho Leung, Joe Ho-Yeung Wong, Siru Lu[1,2]

Hong Kong Monetary Authority, Hong Kong, China

Abstract

The increasing adoption of artificial intelligence (AI) in financial markets has raised concerns about its implications for financial stability. A key concern is that investors using similar AI models may make similar investment decisions and liquidate assets collectively in response to market stress. This collective behaviour could amplify the pro-cyclical price dynamics of asset markets. To empirically examine this concern, we analyse the prospectuses of 87,090 equity open-ended-funds (OEFs) with the assistance of generative AI to identify their applications of AI in investment decision making and construct a matched sample of 2,698 AI-driven OEFs and 2,374 non-AI-driven OEFs with similar characteristics. With this extensive and novel dataset, our analysis suggests that the use of AI could increase the trading correlation among OEFs employing similar yet selective AI-driven investment strategies, compared with their non-AI-driven counterparts. However, these strategies do not necessarily lead to collective asset liquidations during periods of market stress, as they show considerable variation in their responses to market stress and these responses can be either pro-cyclical or counter-cyclical. These findings suggest that the implications of AI usage for financial stability should be assessed with a nuanced approach that acknowledges the heterogeneity of different AI-driven strategies. That said, the pro-cyclical AI-driven strategies identified in this study may warrant closer monitoring.

[1] Leung: vphleung@hkma.gov.hk; Wong: jhywong@hkma.gov.hk; Lu: tslu@hkma.gov.hk

[2] The authors thank Eric Tak-Chuen Wong and participants of the ECB Conference on "The Transformative Power of AI: Economic Implications and Challenges", the EMEAP-BIS Forum on Financial Markets and the HKIMR-IMF-BIS Joint Conference on "Sustaining Financial Stability amid Uncertainty, Fragmentation and Rapid Innovation" for their helpful comments and suggestions. This paper represents the views of the authors, which are not necessarily the views of the Hong Kong Monetary Authority, Hong Kong Academy of Finance Limited, or Hong Kong Institute for Monetary and Financial Research. The above-mentioned entities except the authors take no responsibility for any inaccuracies or omissions contained in the paper.

1. Introduction

Artificial intelligence (AI) refers to the simulation of human intelligence processes by machines, encompassing various subfields, including machine learning (ML), natural language processing (NLP), generative artificial intelligence (GenAI), among others. In recent years, these AI techniques have increasingly permeated nearly every aspect of human life, with financial markets being no exception. With the aid of AI techniques, fund managers can process and analyse large and alternative datasets more efficiently (Milana and Ashta, 2021; Bi et al., 2024; IMF, 2024). This capability enables managers to enhance various investment strategies, including asset value prediction (Bi et al., 2024; Buczynski et al., 2021; Chiang et al., 2021; Daube, 2024; Zhang & Chen, 2017), market sentiment measurement (Xiao & Ke, 2021), thematic screening (Elsaesser et al., 2020), and Environmental, Social and Governance (ESG) risk assessment (Lim, 2024).

Despite the potential benefits documented in the literature, the increasing adoption of AI has raised concerns about its implications for financial stability. A primary concern is whether investors using similar AI models might make analogous investment decisions (Bi et al., 2024; BIS, 2024; Buckmann et al., 2021; Danielsson et al., 2022; Gensler & Bailey, 2020; IMF, 2024; Ramlall, 2018). Specifically, if investors follow similar AI-driven investment instructions, they may collectively liquidate assets at steep discounts in response to market stress. This collective behaviour could amplify the pro-cyclical price dynamics of asset markets. As a result, the application of AI to investment strategies may pose a nascent challenge to financial stability, representing a new area that warrants exploration by financial market regulators.

However, prior studies rarely provide rigorous evidence on this matter. In this paper, we argue that the impacts of AI should not be over-generalised across various applications. Instead, the use of AI may give rise to two distinct scenarios. In a positive scenario, investors use AI to predict market bottoms, informing buy-the-dip strategies during market downturns and thereby mitigating downward pressures on asset markets. Conversely, in a negative

scenario, investors leverage AI to measure market sentiment, facilitating sentiment-driven momentum trading that could amplify downward spirals in asset prices. These different scenarios suggest that the financial stability implications of AI can vary significantly with its application to different investment strategies, which could be either pro-cyclical or counter-cyclical. Therefore, we should avoid evaluating these implications using a one-size-fits-all approach that ignores the heterogeneity of AI-driven strategies.

To examine this issue empirically, we construct a matched sample comprising 2,698 AI-driven equity open-ended funds (OEFs) and 2,374 non-AI-driven OEFs with similar characteristics. This matched sample is selected from a larger dataset of actively-managed equity OEFs, representing approximately 70% of the market aggregate at the end of 2023. Utilising GenAI, we identify the AI adoption of OEFs by analysing 87,090 OEF prospectuses. With this extensive and novel dataset, we investigate three key research questions:

1. *Does the use of AI by OEFs increase the trading correlation among them?*
2. *If yes, does the increased trading correlation result from a similar usage of AI by OEFs, such as applying AI to similar investment strategies?*
3. *Does a similar usage of AI by OEFs lead to collective asset sales during periods of market stress?*

In summary, our findings confirm that AI-driven OEFs exhibit a significantly higher correlation in their trading decisions compared to their non-AI-driven counterparts, but only when they employ similar yet selective AI-driven strategies. However, our analysis reveals that the responses of these AI-driven strategies to market stress vary significantly. This finding suggests that the AI adoption of OEFs does not necessarily lead to collective asset liquidations in periods of market stress. That said, we note that selective AI-driven strategies are pro-cyclical, which may warrant closer monitoring.

This paper contributes to the literature in three key areas. First, to our best knowledge, we provide the first systematic empirical study on the implications of AI application for

financial stability. Importantly, we overcome the challenge of identifying AI usage by analysing a large volume of OEF prospectuses, allowing us to quantify the impacts of AI on asset market dynamics. Second, this study is among the few in the literature to highlight variations across AI-driven investment strategies, offering a more nuanced perspective for market participants, academics, and financial regulators. This is particularly valuable for regulators considering enhanced monitoring of AI usage in financial markets (IMF, 2024). Finally, this study showcases the application of GenAI in empirical financial research, paving the way for future studies utilising this advanced technique.

The remainder of this paper is structured as follows. Section 2 reviews the related literature. Section 3 describes our novel data. Section 4 outlines our hypothesis and empirical models. Section 5 discusses our empirical findings. Section 6 concludes.

2. Literature review

This study relates to four strands of literature. The first strand centres on the application of AI to investment decision-making, a topic that has been extensively explored in previous studies. Notably, Buczynski et al. (2021), Chiang et al. (2021), Daube (2024), Bi et al. (2024) and Zhang and Chen (2017) outline the use of AI in forecasting stock price movements and optimising investment portfolios. Beyond forecasting, a substantial body of literature has examined AI's capability to efficiently analyse lengthy, complex, and multilingual textual information, applying it to various aspects of investment decision-making. For instance, Xiao and Ke (2021) and Beniwal et al. (2024) showcase the use of AI in assessing market sentiment. Elsaesser et al. (2020) demonstrate the application of AI to thematic investing by identifying emerging themes and related companies. Lim (2024) outlines different approaches to assessing ESG risks in investment portfolios with the assistance of AI.

The second strand relates to the implications of AI usage for financial stability, particularly concerns that the use of similar AI techniques by investment funds could lead to a

convergence of their investment decisions. In the event of a significant economic shock, this synchronised response among funds may trigger fire sales, potentially exacerbating pro-cyclicality and market volatility. This concern has been widely documented in the literature, including studies by the BIS (2024), Buckmann et al. (2021), Danielsson et al. (2022), Gensler and Bailey (2020), the IMF (2024), and Ramlall (2018). However, these studies often lack empirical evidence to test their claims. Several studies, including some of those mentioned previously (BIS, 2024; Biais and Woolley, 2011; Buckmann et al., 2021; Danielsson et al., 2022; IMF, 2024; Jain et al., 2016; Malceniece et al., 2019), suggest that the use of similar AI techniques among investors carries similar implications to that of similar trading algorithms from the perspective of financial stability. Nevertheless, to the best of our knowledge, few empirical studies have specifically assessed the implications of AI usage for financial stability.

In addition to the studies that focus on the risks associated with the use of similar AI techniques, some studies examine the risks of AI usage in other aspects. One such risk is whether the use of AI could significantly lower the barriers to algorithmic trading, leveraging its programming capability to facilitate the broader adoption of such trading practices (IMF, 2024). Its increased adoption could potentially propagate the price volatility associated with algorithmic trading (Khandani & Lo, 2011; Zhang, 2010) to the broader financial system. Another risk is the potential for AI to boost short-term leverage among algorithmic traders. As suggested by the IMF (2024), as algorithmic trading becomes more prevalent due to AI, the abnormal returns from exploiting price discrepancies may diminish, incentivising algorithmic traders to adopt higher levels of leverage to enhance their returns. While these issues are pertinent to the implications of AI usage for financial stability, they are not the primary focus of this paper.

The third strand is related to the application of GenAI to empirical studies. As mentioned in the abstract, we employ GenAI to condense a vast volume of OEF prospectuses into precise information regarding their use of AI. This approach has become increasingly

common in empirical studies and is often found to significantly improve accuracy or efficiency compared to manual efforts. For instance, Fatouros et al. (2023) use ChatGPT 3.5 to interpret the sentiment of 2,291 news headlines related to foreign exchange markets as positive, negative, or neutral. They find that the use of GenAI significantly improves sentiment classification accuracy by 35% and the correlation between sentiment and market returns by 36% compared to conventional approaches. Similarly, Hansen and Kazinnik (2023) assess the ability of ChatGPT-4.0 to classify announcements by the Federal Reserve Open Market Committee (FOMC) relative to human assessments. Their analysis demonstrates that ChatGPT-4.0 outperforms human efforts in interpreting the FOMC's language and exhibits a human-like ability to provide explanations for its classifications. Lefort et al. (2024) employ ChatGPT to categorise news articles as positive, negative, or neutral for stock price movements and calculate a market sentiment score. They transform this score into a practical investment strategy for the NASDAQ index and demonstrate that this strategy significantly outperforms the index in terms of Sharpe and Calmar ratios. These studies greatly boost our confidence in selecting GenAI over manual efforts for identifying AI usage by OEFs.

Meanwhile, several related studies have demonstrated the stronger capabilities of relatively advanced versions of GenAI models. Smales (2023) compares ChatGPT-3.5 and ChatGPT-4.0 in classifying the tone of statements issued by the Reserve Bank of Australia and calculating the corresponding sentiment score. This study reveals that ChatGPT-4.0 outperforms its predecessor in comprehending nuances in language. Hansen and Kazinnik (2023) also compare ChatGPT-4.0 with ChatGPT-3.0, showing that the former outperforms in classifying the policy stance of FOMC announcements and in explaining its classifications. Both studies highlight the importance of opting for a relatively advanced GenAI model.

The last strand relates to the use of investment fund prospectuses as a means of gathering qualitative information about fund characteristics, which has been widely applied in academic literature. For instance, Andrikogiannopoulou et al. (2022) measure the commitment of mutual funds to ESG principles by performing a textual analysis on their

prospectuses. Similarly, Birk et al. (2024) categorise an investment fund as sustainable if it explicitly states in its prospectus that it considers sustainability aspects in its investment process. In addition to measuring ESG-related attributes, Kostovetsky and Warner (2020) measure the uniqueness of mutual funds in investment strategies through a textual analysis of these funds' prospectuses. Alda et al. (2022) compile a similar measure for socially responsible mutual funds with reference to these funds' prospectuses. Building on these studies, we will identify AI usage by OEFs primarily by analysing their prospectuses.

3. Data

In this section, we discuss the data and methodology employed in our analysis. Section 3.1 describes our approach to identify the AI usage by OEFs. Section 3.2 outlines how we further identify their various AI-driven investment strategies. Section 3.3 elaborates on the criteria and workflows used to match AI-driven OEFs with their non-AI-driven counterparts that exhibit comparable characteristics. Section 3.4 describes the portfolio data of the matched OEFs used in our analysis. Section 3.5 introduces our measures of market stress.

3.1 Identification of the use of AI

We first compile a large sample of 87,090 equity OEFs from Morningstar Direct[3], representing approximately 70% of the market aggregate at the end of 2023. For these OEFs, we download their prospectuses from the Morningstar Document Library, which typically provides details on their investment strategies and, if any, their application of AI to these strategies. Given the global coverage of OEFs, their prospectuses are written in 27 languages as well as English.[4] English represents 64% of the aggregate, while Korean, German, French, Chinese and Japanese account for most of the rest (Chart 1). To identify the use of AI by each

[3] Morningstar Direct's data providers do not guarantee the accuracy, completeness or timeliness of any information provided by them and shall have no liability for their use.
[4] These languages include Bulgarian, Catalan, Chinese, Croatian, Czech, Danish, Dutch, Estonian, Finnish, French, German, Hungarian, Icelandic, Italian, Japanese, Korean, Latvian, Lithuanian, Malay, Norwegian, Polish, Portuguese, Slovak, Spanish, Swedish, Thai and Turkish.

OEF, we would have had to spend an extensive amount of time analysing such a vast volume of prospectuses individually, requiring exceptional multilingual proficiency, which is definitely a very challenging, if not impossible, task without the aid of GenAI.

To overcome this challenge, we condense the lengthy and complex prospectus of each OEF into precise information regarding its use of AI by inputting the prospectus into GenAI models. These models can rapidly process hundreds of trillions of words and comprehend a variety of languages in a matter of seconds. We select ChatGPT-4o-latest for this process primarily because of its outstanding multilingual proficiency and text evaluation capability[5] compared to its previous generations, including ChatGPT-4 Turbo and ChatGPT-4, as well as other contemporarily advanced GenAI models, including Claude 3 Opus, Gemini Pro 1.5, Gemini Ultra 1.0 and Llama3 400b.

We identify the use of AI with the assistance of ChatGPT-4o-latest and validate the results in three steps. First, we refrain from inputting the entire prospectuses into ChatGPT-4o-latest, as their word count often exceeds the capacity, or context window, of the GenAI model. Instead, we set up a bank of keywords related to various AI subfields in the 28 languages (Table 1 for the English version). We truncate each prospectus to include only the pages containing any of these keywords, along with the immediately previous and subsequent pages, in cases where the keywords appear in the first or last incomplete sentence on that page. Through this process, we detect at least one keyword in the prospectuses of 16,529 OEFs and truncate them for the next step using our own algorithm. Among these prospectuses, English remains the major language, accounting for 74% of the aggregate (Chart 2). The top 5 non-English languages are French, German, Korean, Italian, and Chinese,

[5] See https://help.openai.com/en/articles/8357869-how-to-change-your-language-setting-in-chatgpt for the list of languages supported by ChatGPT; and https://openai.com/index/hello-gpt-4o/ for OpenAI's evaluation on ChatGPT-4o. In addition, we evaluated a multiple of GenAI models available at the time of our study, by requesting these models to identify AI adoption of OEFs in a small sample of OEF prospectuses each of the 28 languages. ChatGPT-4o-latest was one of the few models that could capably identify AI adoption of OEFs with an accurate understanding of all these languages.

capturing nearly all of the rest. Conversely, we presume no use of AI for the remaining 70,561 OEFs.

Of course the occurrence of AI-related keywords in a prospectus does not necessarily imply the application of AI. It is possible that AI-related topics are mentioned in the prospectus as market background or as an investment focus of the OEF. Hence, for each of the truncated prospectuses, we submit a prompt to ChatGPT-4o-latest via Poe, which is a platform that offers access to a wide variety of GenAI-powered models from OpenAI, Anthropic, Google, Meta, and others.[6] For each prompt, we attach a truncated prospectus and ask whether the OEF employs AI techniques to inform its investment decisions.[7] This process results in classifying 5,015 out of 16,529 OEFs as AI users. Additionally, we request that the sentences supporting its classification be extracted from the prospectus for our review.

Lastly, we validate ChatGPT-4o-latest's classifications through two rounds of review. In the first round, we randomly select approximately 10% of the OEFs classified by ChatGPT-4o-latest and manually review whether the GenAI-powered classifications make sense. Our sampling-based validation is consistent with the validation practices of prior studies that utilise GenAI for textual analyses, as discussed in Section 2. Overall, we find that ChatGPT-4o-latest can almost perfectly identify AI adoption by OEFs in this randomly selected sample.

In the second round, we expand the scope of validation beyond the academic literature by examining every OEF identified by ChatGPT-4o-latest. We evaluate its classification by prompting Claude-3.5-Sonnet, which is comparable to ChatGPT-4o-latest in terms of multilingual proficiency and text comprehension.[8] In each prompt, we input ChatGPT-4o-

[6] See https://poe.com/ for Poe.
[7] See Table 2 for the prompt that we submit to ChatGPT-4o-latest. This specifically asks ChatGPT-4o-latest to differentiate between OEFs that use AI for investment decision-making and those that invest in AI-related assets but do not use AI in their investment decision-making process.
[8] Claude-3.5-Sonnet was also one of the few GenAI models that could comprehend the 28 languages and effectively identify the AI adoption of OEFs at the time of our study. Therefore, we employ this comparable GenAI model to validate the outputs of ChatGPT-4o-latest.

latest's classification and the extracted sentences to Claude-3.5-Sonnet and ask whether the classification is reasonable in relation to the sentences.[9] Claude-3.5-Sonnet rejects 248 OEFs classified as AI users and 28 OEFs as non-AI users by ChatGPT-4o-latest. We manually analyse these rejections by Claude-3.5-Sonnet. After manually examining the prospectuses, we reclassify 204 out of the 248 OEFs as non-AI users and 16 out of the 28 OEFs as AI users.[10] Overall, the error rate of ChatGPT-4o-latest is estimated at 1% (i.e. (204 + 16) / 16,529), which is a very low level. We correct for all errors identified in the entire validation process before proceeding to our empirical analysis.

It may be a concern that we use GenAI models that are developed in English-speaking regions, or trained on English-centred data, and therefore may be less able to identify the AI adoption of OEFs which prospectuses are not written in English. However, we note that three quarters of the prospectuses are written in English among the AI-driven OEFs, while the remaining quarter is written in non-English languages. This proportion is remarkably close to the proportion of prospectuses containing AI-related keywords mentioned previously (Chart 2). This finding suggests limited bias towards English prospectuses in our identification of AI adoption by OEFs.

In further analysis, we employ DeepSeek-R1, developed by a China-based AI company, to analyse the prospectuses of AI-driven OEFs. We find that DeepSeek-R1 identifies the same group of OEFs as AI-driven, reaffirming that our identification powered by ChatGPT-4o-latest is not biased towards English prospectuses. Moreover, DeepSeek-R1 is on a par with OpenAI-o1 in terms of its abilities[11], which is an upgraded version of ChatGPT-4o-latest, showcasing the robustness of our identification procedure to using more advanced

[9] See Table 2 for the prompt we create for Claude-3.5-Sonnet to review the classification results generated by ChatGPT-4o-latest.
[10] While Claude-3.5-Sonnet can detect misidentifications made by ChatGPT-4o-latest, this does not necessarily mean that Claude-3.5-Sonnet outperforms ChatGPT-4o-latest, as both of the GenAI models are assigned different tasks. It is possible that ChatGPT-4o-latest detects misidentifications by Claude-3.5-Sonnet, if their tasks are swapped.
[11] See https://github.com/deepseek-ai/DeepSeek-R1 for DeepSeek's evaluation on DeepSeek-R1.

GenAI models. This robustness test is crucial as upgraded versions of ChatGPT-4o-latest and Claude-3.5-Sonnet, such as OpenAI-o1 and Claude-3.7-Sonnet, are now available.

Overall we identify 4,827 OEFs that employ AI to guide their investment decisions, referred to as 'AI-driven OEFs' hereafter. Our novel dataset reveals that the total assets of these OEFs has grown significantly by 45% annually from US$7 billion in 2015 to US$132 billion in 2023, representing approximately 1% of globally actively-managed equity mutual funds (Chart 3).[12] While these AI-driven OEFs are still relatively small in size so far, the potential implications of their AI-driven investment strategies for financial stability may mean that they warrant closer monitoring, given the significant potential for a broader application of these strategies in financial markets in the future.

3.2 Identification of AI-driven investment strategies

Within our sample of AI-driven OEFs, we analyse how they apply AI to their investment strategies. As already mentioned, we have asked ChatGPT-4o-latest to extract sentences indicating the use of AI for each of these OEFs. We further request ChatGPT-4o-latest to digest these sentences and summarise the investment strategies in cases where AI is predominantly used by the OEFs for the following purposes:

a) Asset value prediction (AVP): Applying AI techniques (e.g. ML models) to predict asset values, often to identify undervalued assets for investment.
b) Market sentiment analysis (MSA): Employing AI techniques (e.g. NLP models) to track market sentiment, often used to facilitate sentiment-driven momentum trading strategies.
c) Thematic screening (TS): Leveraging AI techniques (e.g. NLP models) to identify assets aligning with a specific investment theme, with a diversity of

[12] This figure is estimated by comparing the total assets of AI-driven equity OEFs with those of actively-managed equity funds excluding exchange-traded funds sourced from EPFR.

themes being considered by OEFs, including but not limited to sustainability, AI, semiconductors, robotics, and metaverse.

d) ESG risk assessment (ERA): Using AI techniques (e.g. NLP models) to evaluate ESG risks of assets, with a diverse range of risk factors being considered by OEFs, including but not limited to environmental footprint, labour practices, community development, and ethical decision-making.

By referencing the above GenAI-generated summarisation, we give a prompt for each AI-driven OEF to ChatGPT-4o-latest. For each prompt, we attach the sentences extracted from the OEF's prospectus and ask ChatGPT-4o-latest to identify which AI-driven investment strategies are used by the OEF.[13] Following previous studies where textual analyses is powered by GenAI, we randomly select 10% of AI-driven funds and manually review whether the classifications of different AI-driven investment strategies make sense. We find that ChatGPT-4o-latest identifies the various use cases of AI by OEFs in a satisfactory manner within this random sample.

Our analysis shows that 88% of the AI-driven OEFs apply AI to AVP (Chart 4), making it the most popular of AI-driven investment strategies. TS is the second most popular strategy, accounting for 52% of the OEFs. Following that, 15% of the OEFs apply AI to MSA. ERA is the least popular strategy, representing only 3% of the OEFs. These strategies are not mutually exclusive as a given OEF can adopt more than one of them.

In an educated guess, the investment behaviour of these AI-driven investment strategies may differ from one another, particularly in their potential responses to market stress. For instance, OEFs using AI for AVP, referred to as AVP OEFs hereafter, may be better positioned to identify market bottoms and 'buy the dip' by using ML models in times of market stress, potentially mitigating downward pressures on asset markets. Conversely,

[13] See Table 2 for the prompt we create for ChatGPT-4o-latest. This requests ChatGPT-4o-latest to classify each OEF according to its AI usages, which includes (a) AVP, (b) MSA, (c) TS, and (d) ERA, based on the mentions of the OEF's AI usage extracted from their prospectuses. Each OEF may be classified into more than one of these categories.

under the same market conditions, OEFs using AI for MSA, referred to as MSA OEFs, may be relatively able to track the dynamics of market sentiment and facilitate their momentum trading strategies by leveraging NLP models, potentially amplifying the down-cycles of financial markets. Compared to these OEF types, the investment behaviour of the remaining OEFs using AI for TS and ERA, referred to as TS OEFs and ERA OEFs, respectively, may be mostly uncorrelated, since they are focused on a diversity of investment themes or ESG risk factors as mentioned earlier. As such, they may exhibit a mixed response to market stress.

3.3 OEF matching

To identify the effects of AI application on the investment behaviour of OEFs, we need to compare the AI-driven OEFs with their non-AI-driven counterparts. To enable us to identify the effects of AI applications without confounding differences in other OEF-level characteristics, both types of OEFs must possess comparable characteristics.

Procedures for OEF matching are widely documented in the literature and vary significantly across studies. Jin et al. (2022) match bond funds using swing pricing with their counterparts without using swing pricing only if the pairs have the minimum sum of the absolute percentage differences in lagged values of their alphas, total net assets, total expense ratios, ages, average bid-ask spreads of their portfolios, and fractions of shares held by institutional investors. Wu et al. (2024) implement an analogous but relatively lenient approach by minimising the sum of the absolute percentage differences in total net assets, ages, and monthly returns. Becchetti et al. (2015) match socially responsible funds with conventional funds only if the pairs have the closest distance in the coefficients estimated by a five-factor model. Beyond these relatively sophisticated approaches, the following studies employ alternative approaches. Jeffers et al. (2024) match a pair of investment funds on the condition that they are launched in the same year, invest in the same market segments, and have the closest levels of total net assets. Beyond these criteria, Kreander et al. (2005) additionally require the matched pairs to be domiciled in the same jurisdiction. Mallin et al.

(1995) implement a similar approach, but do not require the matched pairs to invest in the same market segment.

Among these studies, we prefer the approach of Jeffers et al. (2024), which strikes a balance between matching stringency and data availability. Following this approach, we match each AI-driven OEF with a non-AI-driven OEF that has the same inception year, the same investment focus, and the closest total net assets. Beyond these requirements, we confine our matched sample to actively-managed OEFs, for the purpose of examining the potential impacts of AI application on the investment behaviours of OEFs. We retrieve our OEF-level variables from Morningstar Direct. Under these criteria, we construct a matched sample of 3,320 AI-driven OEFs and 2,658 non-AI-driven OEFs with replacement.

3.4 Portfolios of OEFs

Within our matched sample, we export data on the matched OEFs' portfolios from Analytics Lab, which is a JupyterLab developed by Morningstar that enables subscribers to programmatically extract portfolio data from Morningstar Direct. The data are monthly spanning the period between January 2022 and June 2024. We select 2022 as the starting point because AI had yet to generate widespread interest before that year, as shown by Google Trends (Chart 5), and might not be used much in the investment fund sector.

We have monthly portfolio data for 2,698 AI-driven OEFs and 2,374 matched non-AI-driven OEFs, which serves as our final sample used for the regression analysis. We exclude OEFs for which there are no monthly portfolio data, since portfolio data at a lower frequency is less meaningful when analysing OEF's portfolio changes. For example, quarterly portfolio data are unable to capture portfolio adjustments taking place within a quarter, merely reflecting the change between two quarter-ends.[14]

[14] OEFs rarely report their portfolio changes on a daily basis. Monthly portfolio data is the highest frequency data that we can use for our empirical analysis.

The portfolio data covers 35,806 assets held by these OEFs in the sample period, with each of these assets which can be held by more than one of these OEFs during the sample period. These assets can be broadly categorised into three asset classes:

a) Cash-like assets: Includes currencies, time deposits, certificates of deposit, money market funds (MMFs), and the US Treasuries.
b) Risky assets: Includes equities, corporate bonds, and non-MMF mutual funds.
c) Derivatives: Includes interest rate derivatives, foreign exchange derivatives, equity derivatives, credit derivatives, commodity derivatives, and index derivatives.

3.5 Measurement of market stress

We use the average of the end-of-day Chicago Board Options Exchange's Volatility Index (VIX) as our primary measure of market stress in a given month. This measure is widely employed in a substantial body of literature, such as Wu et al. (2024), Leung et al. (2024), Converse et al. (2023), Goldstein et al. (2017), and Bao et al. (2011). Following previous research, such as Jin et al. (2022), Kacperczyk et al. (2021) and Rey (2015), we define a month as a stressed period if the VIX for that month exceeds the 75th percentile of the sample. These stressed periods predominantly cover the hawkish pivot by the US FOMC in early 2022 and the subsequent series of significant and rapid interest rate hikes of 50 and 75 basis points during the upward interest rate cycle (Chart 6).

In additional to presenting our regression analysis based on the VIX in Section 4, we find that our results are robust to alternative measures of stressed periods, including the St. Louis Fed Financial Stress Index (SLFFSI) and the US Office of Financial Research Financial Stress Index (OFRFSI), respectively. These measures are also commonly used in the literature to assess global market stress, as shown in Converse et al. (2023) and Ozcelebi (2020).

4. Empirical models

In this section, we detail our empirical models designed to examine the research questions posed in Section 1. We employ an OEF-asset-month level panel dataset, where each row represents a given OEF's holdings of a given asset in any given month. This tri-dimensional panel structure over an OEF-month alternative is preferred because it enables a direct comparison of the matched OEFs in terms of their changes in holdings of identical assets. In contrast, the latter approach would require aggregating holdings of various assets into a single figure for each OEF, which may not be strictly comparable across OEFs due to the typically different compositions of their portfolios. Summary statistics of all variables used in the regression analysis are presented in Table 3.

4.1 Does the use of AI by OEFs increase the trading correlation among them?

We address this question in two steps. First, we measure the trading correlation within AI-driven OEFs and within non-AI-driven OEFs. Second, we test which of these correlations is higher. To do so, we consider the following fixed effects model:

$$\Delta\, log\left(1 + Share_{i,k,t}\right) = \beta_1 \times \Delta\, log\left(1 + \sum_{j \neq i, AI_i = AI_j} Share_{j,k,t}\right) + \beta_2 \times \Delta\, log\left(1 + \sum_{j \neq i, AI_i = AI_j} Share_{j,k,t}\right) \times AI_i + Control_{i,t-1} + \theta_i + \theta_k + \theta_t + \varepsilon_{i,k,t} \quad (1)$$

where $\Delta\, log\left(1 + Share_{i,k,t}\right)$ refers to the logarithmic change in OEF *i*'s holdings of asset *k* in month *t*. AI_i is a dummy variable equal to one if OEF *i* is AI-driven, or zero otherwise. $\Delta\, log\left(1 + \sum_{j \neq i, AI_i = AI_j} Share_{j,k,t}\right)$ represents the logarithmic change in all OEF *j*'s (excluding OEF *i*) holdings of asset *k* in month *t* on the condition that OEF *j*'s usage of AI is the same as that of OEF i (i.e. $AI_i = AI_j$). These holdings are expressed in terms of the number of shares outstanding of an asset, excluding the valuation change. $Control_{i,t-1}$ is a vector of OEF-level control variables, including OEF *i*'s total net assets and age in logarithm,

and monthly return in month *t-1*, denoted as $Size_{i,t-1}$, $Age_{i,t-1}$ and $Ret_{i,t-1}$, respectively. θ_i, θ_k and θ_t denote the fixed effects for OEFs, assets and months. $\varepsilon_{i,k,t}$ is the error term.

Our primary interest lies in the regression coefficients β_1 and β_2 in Equation (1). Specifically, β_1 represents the trading correlation within non-AI driven OEFs, while $\beta_1 + \beta_2$ indicates the trading correlation within AI-driven OEFs. Therefore, β_2 captures the additional trading correlation attributable to the use of AI. If the application of AI by OEFs increases the trading correlation among themselves, we expect β_2 to be significantly positive. To ensure robustness across asset classes, we separately report the estimation of Equation (1) for the sub-samples of cash-like assets, risky assets and derivatives held by OEFs.

It is important to note that the regression coefficient of AI_i is time and asset invariant and is therefore absorbed into the OEF fixed effect θ_i in this set-up, unless we exclude this fixed effect. We prefer to retain θ_i for two reasons. First, the coefficient of AI_i is not our focus, and being absorbed into θ_i does not affect our estimation of β_1 and β_2. Second, a Hausman test (not reported in this paper for brevity) indicates that excluding θ_i from Equation (1) will lead to biased and inconsistent estimates of β_1 and β_2, as θ_i is significantly correlated with the corresponding variables.

4.2 Does the increased trading correlation result from a similar usage of AI by OEFs, such as applying AI to similar investment strategies?

To answer this question, we focus on the sample of AI-driven OEFs and consider a more sophisticated variant of Equation (1):

$$\begin{aligned}\Delta log\left(1 + Share_{i,k,t}\right) = {} & \pi_1 \times \Delta log\left(1 + \textstyle\sum_{j \neq i, AVP_j = 1} Share_{j,k,t}\right) \times AVP_i + \\ & \pi_2 \times \Delta log\left(1 + \textstyle\sum_{j \neq i, AVP_j = 0} Share_{j,k,t}\right) \times AVP_i + \\ & \pi_3 \times \Delta log\left(1 + \textstyle\sum_{j \neq i, MSA_j = 1} Share_{j,k,t}\right) \times MSA_i + \\ & \pi_4 \times \Delta log\left(1 + \textstyle\sum_{j \neq i, MSA_j = 0} Share_{j,k,t}\right) \times MSA_i +\end{aligned}$$

$$\pi_5 \times \Delta log\left(1 + \sum_{j \neq i, OTH_j = 1} Share_{j,k,t}\right) \times OTH_i +$$

$$\pi_6 \times \Delta log\left(1 + \sum_{j \neq i, OTH_j = 0} Share_{j,k,t}\right) \times OTH_i +$$

$$Control_{i,t-1} + \theta_i + \theta_k + \theta_t + \varepsilon_{i,k,t} \quad (2)$$

where AVP_i and MSA_i are dummy variables equal to one if OEF i is an AVP OEF and MSA OEF, respectively, or zero otherwise. OTH_i is a dummy variable equal to one if OEF i is either a TS or ERA OEF, or zero otherwise. They are grouped together because we expect them to exhibit insignificant correlations in their trading decisions as mentioned in Section 3.2. $\Delta log\left(1 + \sum_{j \neq i, AVP_j = 1} Share_{j,k,t}\right)$ refers to the logarithmic change in all AVP OEF j's (excluding OEF i) holdings of asset k in month t. Conversely, $\Delta log\left(1 + \sum_{j \neq i, AVP_j = 0} Share_{j,k,t}\right)$ denotes the logarithmic change in all non-AVP AI-driven OEF j's (excluding OEF i) holdings of asset k in month t. The remaining logarithmic changes are defined similarly, each corresponding to different AI-driven investment strategies as per their respective notions. The control variables and fixed effects are the same as those in Equation (1).

In this set-up, the coefficients of interest are π_1, π_2, π_3, π_4, π_5 and π_6. π_1, π_3 and π_5 which measure the trading correlations within AVP OEFs, MSA OEFs, and TS or ERA OEFs, respectively. Conversely, π_2, π_4 and π_6 measure the trading correlations of AVP OEFs, MSA OEFs, and TS or ERA OEFs, respectively, with their counterparts that adopt different AI-driven investment strategies. If the increased trading correlation is attributable to a similar use of AI-driven investment strategies, we expect the following inequalities to hold: $\pi_1 > \pi_2$, $\pi_3 > \pi_4$ and $\pi_5 > \pi_6$. However, as mentioned in Section 3.2, we only expect a positive trading correlation among either AVP OEFs or MSA OEFs, which implies that only some of these inequalities are satisfied: $\pi_1 > \pi_2$ and $\pi_3 > \pi_4$. Meanwhile, if OEFs using different AI-driven investment strategies invest in an uncorrelated manner, we expect the following inequalities to hold: $\pi_2 \leq 0$, $\pi_4 \leq 0$ and $\pi_6 \leq 0$.

As before, the regression coefficients of AVP_i, MSA_i and OTH_i are time and asset invariant and are therefore absorbed into θ_i in Equation (2), unless we exclude this fixed effect. We retain θ_i for the same reasons given in Section 4.1.

4.3 Does a similar usage of AI by OEFs lead to collective asset sales during periods of market stress?

We address the last question in two steps. First, we investigate whether and how OEFs adjust their holdings of cash-like assets and risky assets in response to market stress.[15] Second, we compare these response of OEFs across different AI-driven investment strategies. To do so, we employ the following fixed effects model:

$$\begin{aligned}\Delta\, log\left(1 + share_{i,k,t}\right) = {} & \beta_1 \times Stress_t \times Cash_k \times NonAI_i + \\ & \beta_2 \times Stress_t \times Risky_k \times NonAI_i + \\ & \pi_1 \times Stress_t \times Cash_k \times AVP_i + \\ & \pi_2 \times Stress_t \times Risky_k \times AVP_i + \\ & \pi_3 \times Stress_t \times Cash_k \times MSA_i + \\ & \pi_4 \times Stress_t \times Risky_k \times MSA_i + \\ & \pi_5 \times Stress_t \times Cash_k \times OTH_i + \\ & \pi_6 \times Stress_t \times Risky_k \times OTH_i + \\ & Control_{i,t-1} + \theta_i + \theta_k + \varepsilon_{i,k,t} \end{aligned} \quad (3)$$

where $NonAI_i$ refers to a dummy variable that equals one if OEF *i* is non-AI-driven, or zero otherwise. $Stress_t$ denotes a dummy variable that equals one if the VIX index in month *t* exceeds the 75th percentile of the sample period, or zero otherwise. $Cash_k$ is a dummy variable that equals one if asset *k* is a cash-like asset, or zero otherwise. $Risky_k$

[15] We conduct a further test using a variant of Equation (3) that incorporate the interaction terms $Stress_t \times Derivative_k \times NonAI_i$, $Stress_t \times Derivative_k \times AVP_i$, $Stress_t \times Derivative_k \times MSA_i$, and $Stress_t \times Derivative_k \times OTH_i$ as the independent variables. $Derivative_k$ is equal to 1 if asset k is a derivative instrument, or 0 otherwise. The results indicate that the estimated coefficients, including β_1, β_2, π_1, π_2, π_3, π_4, π_5, and π_6, are consistent across both equations.

stands for a dummy variable that equals one if asset *k* is a risky asset, or zero otherwise. The control variables in this equation are the same as those in Equations (1) and (2).

In this model, the coefficients of interest are β_1, β_2, π_1, π_2, π_3, π_4, π_5 and π_6, which represent the difference in OEFs' portfolio adjustments during market stress compared to normal times. β_1, π_1, π_3 and π_5 estimate the change in holdings of cash-like assets by non-AI-driven OEFs, AVP OEFs, MSA OEFs, and TS or ERA OEFs in response to market stress, respectively. Conversely, β_2, π_2, π_4 and π_6 estimate the change in holdings of risky assets by non-AI-driven OEFs, AVP OEFs, MSA OEFs, and TS or ERA OEFs in response to market stress, respectively. As discussed in Section 3.2, we expect the investment cyclicality of OEFs to vary significantly across different AI-driven investment strategies. AVP OEFs are expected to behave in a counter-cyclical manner, using their cash buffer to buy the dip during periods of market stress. Therefore, we anticipate that π_1 will be significantly negative and π_2 will be significantly positive. In contrast, MSA OEFs are expected to be pro-cyclical, shifting from risky assets to cash during periods of market stress. Consequently, we expect π_3 to be significantly positive and π_4 to be significantly negative. The remaining TS or ERA OEFs are expected to behave in a non-cyclical way and therefore we expect π_5 and π_6 to be insignificant. Furthermore, if we find that MSA OEFs liquidate risky assets in response to market stress, we will test whether their investment behaviour is more pro-cyclical than that of non-AI-driven OEFs by performing a Wald test on the inequality $\pi_4 < \beta_2$, where $\beta_2 < 0$ if non-AI-driven OEFs also exhibit pro-cyclical investment behaviour.

Before we report our empirical results, there are three technical issues to mention. First, we exclude derivatives held by OEFs due to a lack of information on the risk profiles of these derivatives, which could be used for speculation or hedging. However, our results remain robust to including changes in OEFs' holdings of these derivatives in Equation (3). Second, we are unable to control for unobservable time variation θ_t, as this will result in the absorption of our coefficients of interest into θ_t. Third, several interaction terms, including $Cash_k \times NonAI_i$, $Risky_k \times NonAI_i$ (equivalently the disaggregation of $NonAI_i$), $Cash_k \times$

AVP_i, $Risky_k \times AVP_i$ (equivalently the disaggregation of AVP_i), $Cash_k \times MSA_i$, $Risky_k \times MSA_i$ (equivalently the disaggregation of MSA_i), $Cash_k \times OTH_i$, and $Risky_k \times OTH_i$ (equivalently the disaggregation of OTH_i),[16] are time-invariant and are absorbed into the fixed effects θ_i and θ_k, unless we exclude one of these effects. Nevertheless, we retain both fixed effects for the same reasons as mentioned in Sections 4.1 and 4.2.

5. Empirical findings

In estimating Equations (1) – (3) in this section, the standard error of the coefficients is clustered at the level of OEFs and assets to account for the possibility that observations within each cluster are not independently and identically distributed. We refrain from clustering at the time level, as there are only 30 months in our panel data which barely meets the rule of thumb for valid clustering (Cameron and Miller, 2015). Clustering at the time level in panel data with a short sample frame can lead to underestimated standard errors and an increased likelihood of over-rejecting the null hypotheses.

5.1 Does the use of AI by OEFs increase the trading correlation among them?

Our empirical findings suggest that the use of AI by OEFs can increase the trading correlation among them. Specifically, the results from Equation (1) show that the estimated β_1 and β_2 are 0.14 and 0.04 at a 1% level of statistical significance, respectively (Column 1 of Table 4). These figures indicate that the trading correlation among non-AI-driven OEFs is positive at 0.14, while the trading correlation among AI-driven OEFs is also positive at 0.18 (i.e. 0.14 + 0.04). In other words, other things being equal, the use of AI significantly increases the trading correlation among OEFs by 0.04. This estimated effect is also economically significant as it implies a notable increase in trading correlations of approximately one-third (i.e. 0.04 / 0.14).

[16] Similarly, these interaction terms can be interpreted as the disaggregation of $Cash_k$ into $Cash_k \times NonAI_i$, $Cash_k \times AVP_i$, $Cash_k \times MSA_i$, and $Cash_k \times OTH_i$, and the disaggregation of $Risky_k$ into $Risky_k \times NonAI_i$, $Risky_k \times AVP_i$, $Risky_k \times MSA_i$, and $Risky_k \times OTH_i$.

We also find that the use of AI increases the trading correlation of OEFs across various asset classes. In the sub-sample of cash-like assets held by OEFs, the estimated β_1 and β_2 are 0.09 and 0.01 at a 1% level of statistical significance, respectively (Column 2 of Table 4). In the sub-sample of risky assets, the estimated β_1 and β_2 are 0.23 and 0.03 at a 1% level of statistical significance, respectively (Column 3 of Table 4). In the sub-sample of derivatives, the estimated β_1 and β_2 are 0.12 and 0.08 at a 1% level of statistical significance, respectively (Column 4 of Table 4).

5.2 Does the increased trading correlation result from a similar usage of AI by OEFs, such as applying AI to similar investment strategies?

We find that the use of AI by OEFs increases the trading correlation among them only when they apply AI to identical investment strategies. Specifically, the results of Equation (2) reveal that the estimated π_1 and π_3 are 0.20 and 0.13 at a 1% level of statistical significance, respectively (Column 1, Table 5). Conversely, the estimated π_2 and π_4 are -0.03 and -0.11 at a 1% of statistical significance, respectively (Column 1, Table 5). These figures indicate that both types of OEFs exhibit a positive trading correlation only with other OEFs that apply AI to the same investment strategy, while demonstrating a negative trading correlation with other OEFs that apply AI to different strategies. Wald tests show that π_1 and π_3 are significantly greater than π_2 and π_4 at a 1% level of statistical significance, respectively (Column 1, Table 5). Meanwhile, the remaining π_5 and π_6 are estimated to be statistically insignificant (Column 1, Table 5), which aligns with our prior expectation that the investment behaviours of either TS or ERA OEFs are uncorrelated. Taken together, our analysis suggests that the use of AI by OEFs does not necessarily increase the trading correlation among them, and those positive correlations are confined to those OEFs that share similar yet selective uses of AI in their trading strategies, including AVP and MSA.

The estimation results of Equation (2) are robust across various asset classes. In the sub-sample of cash-like assets held by OEFs, the estimated π_1 and π_3 are significantly

positive at 0.08 and 0.09, while the estimated π_2 and π_4 are significantly negative at -0.02 and -0.04, respectively (Column 2, Table 5). In the sub-sample of risky assets, the estimated π_1 and π_3 are significantly positive at 0.19 and 0.18, which are significantly larger than the estimated π_2 and π_4 at 0.01 and -0.14, respectively (Column 3, Table 5). In the sub-sample of derivatives, we estimate significantly positive values for π_1 and π_3 at 0.25 and 0.14, compared to the estimated negative values for π_2 and π_4 at -0.04 and -0.15 (Column 4, Table 5). The remaining π_5 and π_6 are estimated to be insignificant in these sub-samples.

5.3 Does a similar usage of AI by OEFs lead to collective asset sales during periods of market stress?

Our empirical results indicate that a similar usage of AI by OEFs does not necessarily lead to collective asset sales during periods of market stress, as their responses to market stress vary significantly with different investment strategies. Specifically, the estimation results of Equation (3) shows that the estimates of π_1 and π_2 are -0.03 and 0.02 at a 1% level of statistical significance, respectively (Column 1, Table 6). This implies that AVP OEFs invest counter-cyclically by using their cash buffer to buy the dip in times of market stress. Conversely, the estimates of π_3 and π_4 are at 0.04 and -0.11, respectively (Column 1, Table 6), suggesting that MSA OEFs invest pro-cyclically by substituting risky assets for cash-like assets in times of market stress. The estimates of π_5 and π_6 are statistically insignificant, aligning with our hypothesis that TS or ERA OEFs respond to market stress in a rather mixed manner (Column 1, Table 6).

Furthermore, our findings suggest that the investment behaviour of MSA OEFs is more pro-cyclical than non-AI-driven OEFs. Specifically, the estimate of π_4 of -0.11 is more negative than the estimate of β_2 of -0.02 (Column 1, Table 6), and this difference is statistically significant at the 1% level. The difference is also economically significant, as the magnitude of the estimated π_4 is about 5 times that of the estimated β_2 (i.e. -0.11/-0.02). This

finding underscores the potential for AI, albeit limited to selective applications, to exacerbate the pro-cyclicality of OEFs' investment behaviour.

As mentioned in Section 3.5, we use two variants of the dummy variable $Stress_t$ to assess the robustness of our results to different measures of market stress. Both variants are set to one if the SLFFSI and OFRFSI in month *t* exceed the 75th percentile of the sample period, respectively, or zero otherwise. The robustness tests show that the estimation results from Equation (3) remain largely consistent across these alternative measures of market stress (Columns 2 and 3, Table 6).

6. Conclusion and implications

Taken together, the rise of AI yields a wide range of potential benefits across virtually every aspect of human life. This paper provides one example in the efficiency of GenAI in analysing a vast volume of textual information, far exceeding the capabilities of manual review. Despite substantial potential benefits from the use of AI by financial markets, regulators still need to remain vigilant about possible increases in material risks from the increased use of AI in investment decision-making.

Unlike a substantial body of literature that suggests the impacts of AI on financial markets are destabilising albeit with scant empirical evidence, this paper pioneers a systemic empirical study to examine the implications of AI application for financial stability based on a large panel data set of OEFs. Our analysis shows that the use of AI by OEFs could increase the trading correlation among them, relative to their non-AI-driven counterparts, but only when OEFs apply AI to similar yet selective investment strategies. However, the adoption of similar AI-driven strategies does not necessarily lead to collective asset liquidations during periods of market stress, as their responses to market stress exhibit significant variation.

From the perspective of financial stability, these findings carry two policy implications. First, the implications of AI usage for financial stability should be evaluated using a nuanced approach that acknowledges the heterogeneity of AI-driven investment

strategies, which exhibit significant variation in their responses to market stress. Second, the pro-cyclical AI-driven investment strategies identified in this study warrants particularly closer monitoring, given the potential risks to financial sector stability.

Lastly, it is worth mentioning two caveats to this study. First, our inferences regarding the use of AI are primarily based on the information extracted from OEFs' prospectuses. OEFs may not fully disclose their use of AI, and even when they do, their AI-driven investment strategies might still involve a degree of manual judgement (e.g. a human-in-the-loop approach). This may pose certain challenges to accurately identifying the impacts of AI adoption. Second, our assessment of the implications of AI usage for financial stability does not cover the effects of AI on other aspects of investment behaviour, such as high-frequency trading or leverage. Investigating these aspects requires different data and therefore, readers should be careful in interpreting our results.

References

Alda, M., Muñoz, F., & Vargas, M. (2022). Product differentiation in the socially responsible mutual fund industry. *Journal of Multinational Financial Management*, *64*, 100730.

Andrikogiannopoulou, A., Krueger, P., Mitali, S., & Papakonstantinou, F. (2023). Discretionary information in ESG investing: A text analysis of mutual fund prospectuses.

Bao, J., Pan, J., & Wang, J. (2011). The illiquidity of corporate bonds. *The Journal of Finance*, 66(3), 911-946.

Bank for International Settlements (BIS). (2024). BIS Annual Economic Report 2024: Artificial intelligence and the economy: implications for central banks.

Becchetti, L., Ciciretti, R., Dalò, A., & Herzel, S. (2015). Socially responsible and conventional investment funds: performance comparison and the global financial crisis. *Applied Economics*, *47*(25), 2541-2562.

Beniwal, M., Singh, A., & Kumar, N. (2024). Forecasting multistep daily stock prices for long-term investment decisions: A study of deep learning models on global indices. *Engineering Applications of Artificial Intelligence*, *129*, 107617.

Biais, B., & Woolley, P. (2011). High frequency trading. *Manuscript, Toulouse University, IDEI*.

Bi, S., Bao, W., Xiao, J., Wang, J., & Deng, T. (2024). Application and practice of AI technology in quantitative investment. *arXiv preprint arXiv:2404.18184*.

Birk, K., Jacob, S., & Wilkens, M. (2024). What attracts sustainable fund flows? Prospectus vs. ratings.

Buckmann, M., Haldane, A., & Hüser, A. C. (2021). Comparing minds and machines: implications for financial stability. *Oxford Review of Economic Policy*, *37*(3), 479-508.

Buczynski, W., Cuzzolin, F., & Sahakian, B. (2021). A review of machine learning experiments in equity investment decision-making: why most published research findings do not live up to their promise in real life. *International Journal of Data Science and Analytics*, *11*, 221-242.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, *50*(2), 317-372.

Chiang, D. L., Wang, S. K., Lin, Y. N., Yang, C. Y., Shen, V. R., Juang, T. T. Y., & Liao, T. Y. (2021). Development and evaluation of a novel investment decision system in cryptocurrency market. *Applied Artificial Intelligence*, *35*(14), 1169-1195.

Converse, N., Levy-Yeyati, E., & Williams, T. (2023). How ETFs amplify the global financial cycle in emerging markets. *The Review of Financial Studies*, *36*(9), 3423-3462.

Danielsson, J., Macrae, R., & Uthemann, A. (2022). Artificial intelligence and systemic risk. *Journal of Banking & Finance*, *140*, 106290.

Daube, C. H. (2024). Artificial intelligence in financial and investment decision-making (No. 2/2024). Working Papers des IUCF.

Elsaesser, G., Gardin, F., Kolrep, M., & Rosentritt, M. (2020). Thematic Innovation Investing with Textual Data Analysis. *Risk & Reward*, 20-26.

Fatouros, G., Soldatos, J., Kouroumali, K., Makridis, G., & Kyriazis, D. (2023). Transforming sentiment analysis in the financial domain with ChatGPT. *Machine Learning with Applications*, *14*, 100508.

Gensler, G., & Bailey, L. (2020). Deep learning and financial stability. *Available at SSRN 3723132*.

Goldstein, I., Jiang, H., & Ng, D. T. (2017). Investor flows and fragility in corporate bond funds. *Journal of Financial Economics*, *126*(3), 592-613.

Hansen, A. L., & Kazinnik, S. (2023). Can chatgpt decipher fedspeak. *Available at SSRN*.

International Monetary Fund (IMF). (2024). Global financial stability report: Steadying the course: Uncertainty, artificial intelligence, and financial stability. Washington, DC: International Monetary Fund.

Jain, P. K., Jain, P., & McInish, T. H. (2016). Does high-frequency trading increase systemic risk?. *Journal of Financial Markets*, *31*, 1-24.

Jeffers, J., Lyu, T., & Posenau, K. (2024). The risk and return of impact investing funds. *Journal of Financial Economics*, *161*, 103928.

Jin, D., Kacperczyk, M., Kahraman, B., & Suntheim, F. (2022). Swing pricing and fragility in open-end mutual funds. *The Review of Financial Studies*, *35*(1), 1-50.

Kacperczyk, M. T., Pérignon, C., & Vuillemey, G. (2021). The private production of safe assets. *Journal of Finance*, 76(2), 495-535.

Khandani, A. E., & Lo, A. W. (2011). What happened to the quants in August 2007? Evidence from factors and transactions data. *Journal of Financial Markets*, 14(1), 1-46.

Kostovetsky, L. and Warner, J.B. (2020), Measuring Innovation and Product Differentiation: Evidence from Mutual Funds. *The Journal of Finance*, 75: 779-823.

Kreander, N., Gray, R. H., Power, D. M., & Sinclair, C. D. (2005). Evaluating the performance of ethical and non-ethical funds: a matched pair analysis. *Journal of Business Finance & Accounting*, *32*(7-8), 1465-1493.

Lefort, B., Benhamou, E., Ohana, J. J., Saltiel, D., Guez, B., & Jacquot, T. (2024). Sentiment Analysis of Bloomberg Markets Wrap Using ChatGPT: Application to the NASDAQ. *Available at SSRN 4780150*.

Leung, D., Wong, J., & Fong, T. P. W. (2024). Run risks of cash-redeemable ETFs. *Pacific-Basin Finance Journal*, 85, 102356.

Lim, T. (2024). Environmental, social, and governance (ESG) and artificial intelligence in finance: State-of-the-art and research takeaways. *Artificial Intelligence Review*, 57, 76.

Malceniece, L., Malcenieks, K., & Putniņš, T. J. (2019). High frequency trading and co-movement in financial markets. *Journal of Financial Economics*, *134*(2), 381-399.

Mallin, C. A., & Saadouni, B. (1995). The financial performance of ethical investment funds. *Journal of Business Finance & Accounting*, *22*(4).

Milana, C., & Ashta, A. (2021). Artificial intelligence techniques in finance and financial markets: a survey of the literature. *Strategic Change*, *30*(3), 189-209.

Ozcelebi, O. (2020). Assessing the impacts of financial stress index of developed countries on the exchange market pressure index of emerging countries. *International Review of Economics & Finance*, *70*, 288-302.

Ramlall, I. (2018). FinTech and the financial stability board. In *Understanding Financial Stability* (pp. 71-81). Emerald Publishing Limited.

Rey, H. (2015). Dilemma not trilemma: The global financial cycle and monetary policy independence (NBER Working Paper No. 21162). National Bureau of Economic Research.

Smales, L. A. (2023). Classification of RBA monetary policy announcements using ChatGPT. *Finance Research Letters*, *58*, 104514.

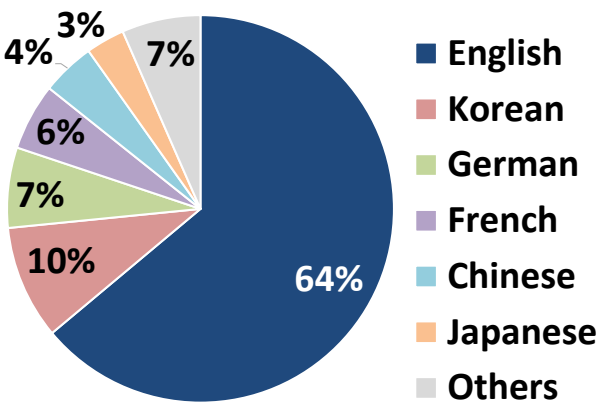
Wu, G., Wong, J., & Fong, T. (2024). Does swing pricing reduce investment funds' liquidity risk in times of market stress? – Evidence from the March-2020 episode. *North American Journal of Economics and Finance*, 72, 102118.

Xiao, F., & Ke, J. (2021). Pricing, management and decision-making of financial markets with artificial intelligence: introduction to the issue. *Financial Innovation*, *7*, 1-3.

Zhang, F. (2010). High-frequency trading, stock volatility, and price discovery. *Available at SSRN 1691679*.

Zhang, X., & Chen, Y. (2017). An artificial intelligence application in portfolio management. In *International Conference on Transformations and Innovations in Management (ICTIM 2017)* (pp. 775-793). Atlantis Press.

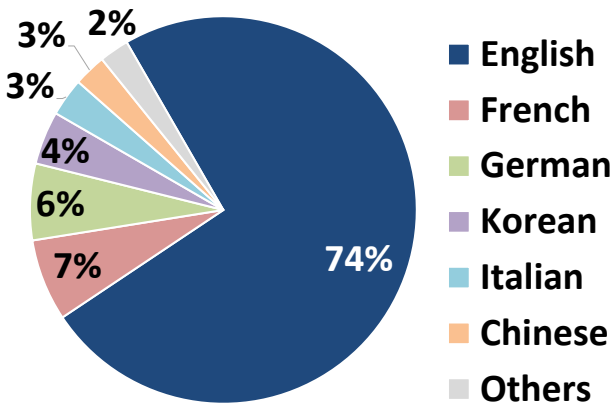
Chart 1. Share of languages across OEF prospectuses



Note: The pie represents the proportion of languages in use across all OEF prospectuses in our sample.

Sources: Morningstar Direct, EPFR and authors' estimation.

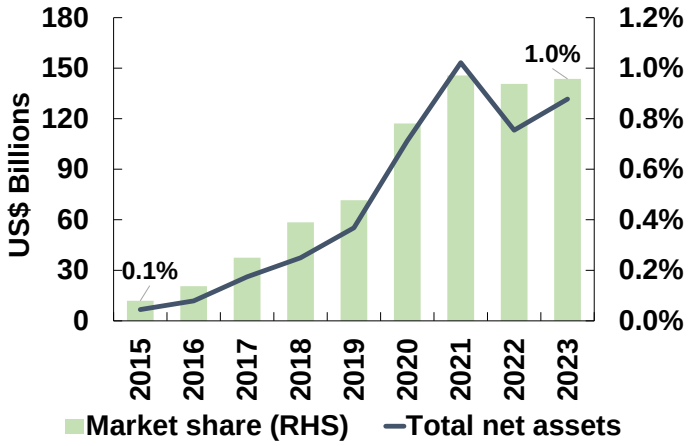
Chart 2. Share of languages across OEF prospectuses containing AI-related keywords



Note: The pie represents the proportion of languages in use across all OEF prospectuses containing at least one AI-related keyword in our sample.

Sources: Morningstar Direct and authors' estimation.

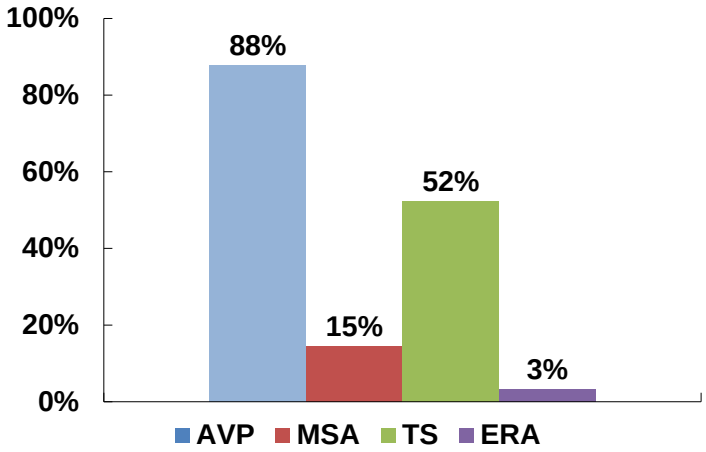
Chart 3. Total assets and market share of AI-driven equity OEFs



Note: The curve represents the total net assets of AI-driven equity OEFs in our sample. Each bar indicates the share of these OEFs in the global universe of actively-managed equity mutual funds in any given year.

Sources: Morningstar Direct, EPFR and authors' estimation.

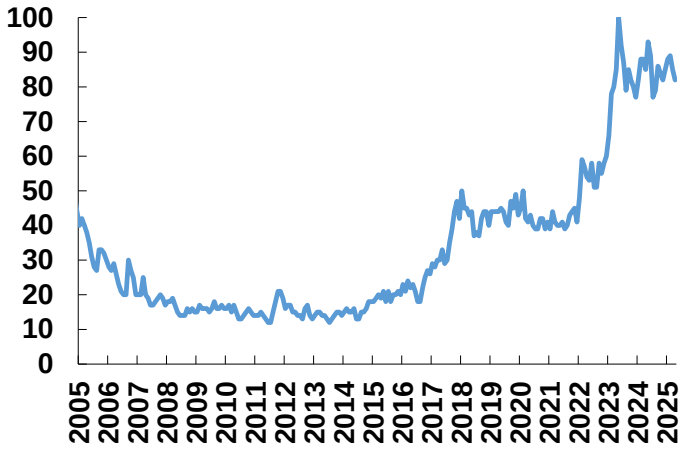
Chart 4. Share of AI-driven investment strategies adopted by AI-driven equity OEFs



Note: Each bar represents the share of AI-driven OEFs adopting each type of AI-driven investment strategy at the end of 2023, in terms of total net assets. The sum of the bars exceeds one as OEFs can adopt more than one type of AI-driven investment strategy.

Sources: Morningstar Direct and authors' estimation.

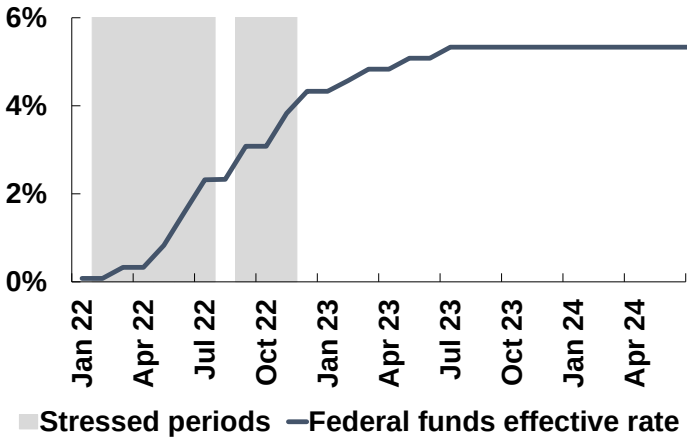
Chart 5. Google Trends of keyword "artificial intelligence"



Note: The curve indicates Google Trends of the keyword "artificial intelligence" from 2005 to 2025 reflecting public interest in this keyword.

Source: Google.

Chart 6. The US federal funds effective rate and periods of market stress



Note: The curve indicates the US federal funds effective rate, with the grey areas representing periods of market stress identified in this study based on the level of the VIX index.

Sources: Federal Reserve Economic Data and authors' estimation.

Table 1. Keywords related to AI (English version)

artificial intelligence	supervised learning
AI	semi-supervised learning
machine learning	cluster analysis
deep learning	swarm
decision tree	sentiment analysis
neutral network	textual model
natural language processing	text mining
reinforcement learning	attention mechanism
unsupervised learning	cognitive computing

Table 2. Prompts made to GenAI models

GenAI model:	ChatGPT-4o-latest
Objective:	To identify the use of AI by each OEF in its investment decision-making
Prompt:	The attachment is a prospectus of an investment fund. According to the content, please identify whether the investment fund (1) uses artificial intelligence (AI) for making investment decision or (2) invests in AI-related securities or (3) both or (4) neither. For the first three circumstances, please cite the relevant sentences as evidence in the next lines.
Attachment:	An OEF's truncated prospectus

Table 2. Prompts made to GenAI models (continued)

GenAI model:	Claude-3.5-Sonnet
Objective:	To validate the classification by ChatGPT-4o-latest for each OEF
Prompt:	ChatGPT-4o-latest has classified an investment fund as AI-driven/non-AI-driven fund based on the fund's prospectus, and extract sentences from this prospectus to support its classification. The attachment contains these sentences. Please examine whether ChatGPT-4o-latest classifies this fund correctly based on these sentences and explain.
Attachment:	A file containing sentences extracted from an OEF's truncated prospectus by ChatGPT-4o-latest

GenAI model:	ChatGPT-4o-latest
Objective:	To identify the use case(s) of AI by AI-driven OEFs
Prompt:	The attachment contains sentences describing how an investment fund applies AI to its investment strategies. According to the content, please classify the use case(s) of AI by this fund. These use case(s) can be (1) using AI techniques to predict asset values to inform investment decisions; (2) using AI techniques to measure market sentiment to inform investment decisions; (3) using AI techniques to identify assets aligning with specific investment themes; (4) using AI techniques to measure ESG risks of assets.
Attachment:	A file containing sentences extracted from an OEF's truncated prospectus by ChatGPT-4o-latest

Table 3. Summary statistics

	Mean	SD	25pt	50pt	75pt	Obs.
$\Delta\, log(1 + Share_{i,k,t})$	-0.03	4.27	0.00	0.00	0.00	41,601,089
$\Delta\, log\left(1 + \sum_{j \neq i, AI_i = AI_j} Share_{j,k,t}\right)$	-0.07	2.96	-0.39	0.00	0.27	41,601,089
AI_i	0.55	0.50	0.00	1.00	1.00	41,601,089
$NonAI_i$	0.45	0.50	0.00	0.00	1.00	41,601,089
$Size_{i,t-1}$	14.41	3.44	11.92	15.07	17.13	41,601,089
$Age_{i,t-1}$	1.44	0.54	1.07	1.50	1.87	41,601,089
$Ret_{i,t-1}$	0.04	6.25	-3.77	0.04	4.01	41,601,089
AVP_i	0.82	0.39	1.00	1.00	1.00	21,819,354
MSA_i	0.33	0.47	0.00	0.00	1.00	21,819,354
OTH_i	0.49	0.50	0.00	0.00	1.00	21,819,354
$\Delta\, log\left(1 + \sum_{j \neq i, AVP_j = 1} Share_{j,k,t}\right)$	-0.09	3.36	-0.53	0.00	0.37	21,819,354
$\Delta\, log\left(1 + \sum_{j \neq i, AVP_j = 0} Share_{j,k,t}\right)$	-0.04	4.06	-0.19	0.00	0.20	21,819,354
$\Delta\, log\left(1 + \sum_{j \neq i, MSA_j = 1} Share_{j,k,t}\right)$	-0.06	3.45	-0.38	0.00	0.30	21,819,354
$\Delta\, log\left(1 + \sum_{j \neq i, MSA_j = 0} Share_{j,k,t}\right)$	-0.06	3.27	-0.50	0.00	0.38	21,819,354
$\Delta\, log\left(1 + \sum_{j \neq i, OTH_j = 1} Share_{j,k,t}\right)$	-0.04	3.67	-0.37	0.00	0.31	21,819,354
$\Delta\, log\left(1 + \sum_{j \neq i, OTH_j = 0} Share_{j,k,t}\right)$	-0.12	3.48	-0.56	0.00	0.37	21,819,354
$Stress_t$	0.25	0.43	0.00	0.00	1.00	29,829,902
$Cash_k$	0.28	0.45	0.00	0.00	1.00	29,829,902
$Risky_k$	0.72	0.45	0.00	1.00	1.00	29,829,902

Table 4. Estimation results of Equation (1)

Matched OEFs: of which	$\Delta log(1 + Share_{i,k,t})$			
	All assets	Cash-like assets	Risky assets	Derivatives
	(1)	(2)	(3)	(4)
$\Delta log\left(1 + \sum_{j \neq i, AI_i = AI_j} Share_{j,k,t}\right)$	0.14***	0.09***	0.23***	0.12***
$\Delta log\left(1 + \sum_{j \neq i, AI_i = AI_j} Share_{j,k,t}\right) \times AI_i$	0.04***	0.01***	0.03***	0.08***
$Size_{i,t-1}$	0.001**	-0.00	0.00***	-0.00***
$Age_{i,t-1}$	-0.11***	-0.05***	-0.06***	-0.31***
$Ret_{i,t-1}$	0.00***	0.00***	0.00***	0.01***
Wald test: $\beta_1 + \beta_2$	0.18***	0.10***	0.26***	0.20***
Fixed effects:				
OEFs	Yes	Yes	Yes	Yes
Assets	Yes	Yes	Yes	Yes
Months	Yes	Yes	Yes	Yes
No. of OEFs	5,072	4,993	5,072	3,601
No. of assets	35,806	4,124	25,953	5,729
No. of months	30	30	30	30
No. of observations	41,601,089	8,238,709	21,591,193	11,771,187

Note: ***, ** and * represent a 1%, 5% and 10% level of statistical significance. The standard error of coefficients is clustered at the level of OEFs and assets.

Table 5. Estimation results of Equation (2)

Matched AI-driven OEFs: of which	$\Delta log(1 + Share_{i,k,t})$			
	All assets	Cash-like assets	Risky assets	Derivatives
	(1)	(2)	(3)	(4)
$\Delta log\left(1 + \sum_{j \neq i, AVP_j = 1} Share_{j,k,t}\right) \times AVP_i$	0.20***	0.08***	0.19***	0.25***
$\Delta log\left(1 + \sum_{j \neq i, AVP_j = 0} Share_{j,k,t}\right) \times AVP_i$	-0.03***	-0.02***	0.01***	-0.04***
$\Delta log\left(1 + \sum_{j \neq i, MSA_j = 1} Share_{j,k,t}\right) \times MSA_i$	0.13***	0.09***	0.18***	0.14***
$\Delta log\left(1 + \sum_{j \neq i, MSA_j = 0} Share_{j,k,t}\right) \times MSA_i$	-0.11***	-0.04***	-0.14***	-0.15***
$\Delta log\left(1 + \sum_{j \neq i, OTH_j = 1} Share_{j,k,t}\right) \times OTH_i$	0.02	0.01	0.03	0.02
$\Delta log\left(1 + \sum_{j \neq i, OTH_j = 0} Share_{j,k,t}\right) \times OTH_i$	-0.01	-0.01	-0.04	-0.01
$Size_{i,t-1}$	0.00	0.00**	0.00***	-0.01***
$Age_{i,t-1}$	-0.07***	-0.05***	0.01*	-0.26***
$Ret_{i,t-1}$	0.00***	0.01***	0.00***	0.01***
Wald test: $\pi_1 - \pi_2$	0.23***	0.10***	0.18***	0.29***
Wald test: $\pi_3 - \pi_4$	0.24***	0.13***	0.32***	0.29***
Wald test: $\pi_5 - \pi_6$	0.03	0.02	0.07	0.03
Fixed effects:				
OEFs	Yes	Yes	Yes	Yes
Assets	Yes	Yes	Yes	Yes
Months	Yes	Yes	Yes	Yes
No. of OEFs	2,698	2,692	2,698	2,106
No. of assets	14,473	2,232	11,476	765
No. of months	30	30	30	30
No. of observations	21,819,354	4,991,875	10,867,962	5,959,517

Note: ***, ** and * represent a 1%, 5% and 10% level of statistical significance. The standard error of coefficients is clustered at the level of OEFs and assets.

Table 6. Estimation results of Equation (3)

$Stress_t$ defined based on	$\Delta log(1 + share_{i,k,t})$		
	VIX	SLFFSI	OFRFSI
	(1)	(2)	(3)
$Stress_t \times Cash_k \times NonAI_i$	-0.06***	0.04***	-0.07***
$Stress_t \times Risky_k \times NonAI_i$	-0.02***	0.08***	0.01***
$Stress_t \times Cash_k \times AVP_i$	-0.03***	-0.02***	-0.09***
$Stress_t \times Risky_k \times AVP_i$	0.02***	0.09***	0.06***
$Stress_t \times Cash_k \times MSA_i$	0.04***	0.00	0.00
$Stress_t \times Risky_k \times MSA_i$	-0.11***	-0.11***	-0.11***
$Stress_t \times Cash_k \times OTH_i$	-0.01	0.01	-0.01
$Stress_t \times Risky_k \times OTH_i$	-0.01	-0.01	-0.01
$Size_{i,t-1}$	0.00***	0.00***	0.00***
$Age_{i,t-1}$	-0.36***	-0.26***	-0.34***
$Ret_{i,t-1}$	0.00***	0.00***	0.00***
Wald test: $\pi_4 - \beta_2$	-0.09***	-0.19***	-0.12***
Fixed effects:			
OEFs	Yes	Yes	Yes
Assets	Yes	Yes	Yes
Months	No	No	No
No. of OEFs	5,072	5,072	5,072
No. of assets	30,077	30,077	30,077
No. of months	30	30	30
No. of observations	29,829,902	29,829,902	29,829,902

Note: ***, ** and * represent a 1%, 5% and 10% level of statistical significance. The standard error of coefficients is clustered at the level of OEFs and assets.