ChatGPT for (Finance) Research: The Bananarama Conjecture

Michael Dowling*[1] and Brian Lucey[2,3,4,5]

[1]DCU Business School, Dublin City University, Ireland
[2]Trinity Business School, Trinity College Dublin, Ireland
[3]University of Abu Dhabi, Zayed City, Abu Dhabi
[4]University of Economics, Ho Chi Minh City, Vietnam
[5]Jiangxi University of Economics and Finance, Nanchang, China

Abstract

We show, based on ratings by finance journal reviewers of generated output, that te recently released AI chatbot ChatGPT can significantly assist with finance research. In principle, these results should be generalisable across research domains. There are clear advantages for idea generation and data identification. The technology, however, is weaker on literature synthesis and developing appropriate testing frameworks. Importantly, we further demonstrate that the extent of private data and researcher domain expertise input, are key factors in determining the quality of output. We conclude by considering the implications, particularly the ethical implications, which arise from this new technology.

JEL codes: G00; G10
Keywords: ChatGPT; artificial intelligence; finance research; ethics

*Michael Dowling: DCU Business School, Dublin City University, Glasnevin, Dublin 9, Ireland. Email: michael.dowling@dcu.ie

1 Introduction

"It ain't what you do, it's the way that you do it

And that's what gets results"

— Song lyrics by Bananarama and Fun Boy Three (1982)

ChatGPT is an artificial intelligence language model introduced in November 2022 providing generated conversational responses to question prompts. The model is trained with a blend of reinforcement learning algorithms and human input on over 150 billion parameters[1]. The platform reached a million users in just its first week open to the public and has been quickly coined "the industry's next big disrupter" (Grant and Metz, 2022) due to the perceived quality of response output from the model.

One early academic study found the platform capable of passing the notoriously-complex common core of US professional legal accreditation examinations (Bommarito II and Katz, 2022). Another author managed to produce a reasonably-comprehensive guide to quantitative trading, almost exclusively through *ChatGPT* output (Marti, 2022). A range of professions even set themselves to existential pondering as to whether they have suddenly been made redundant; including educators (Herman, 2022), lawyers (Greene, 2022), and, to cover as many worried professional bases as possible, 'all writers' (Warner, 2023). It's quite the entrance for the new technology.

We are interested in the extent to which *ChatGPT* can assist with the production of research studies; in this case, finance research. Initial research has explored some limited aspects of this question. A broad perspective on the emergent role for AI

[1] https://openai.com/blog/chatgpt/

in the production of scientific research is taken by Grimaldi and Ehrler (2023) and Hutson et al. (2022). While Alshater (2022) suggests that ChatGPT *should* be useful for a range of tasks involved in constructing a research study, but without empirical testing.

Most of the applied research focuses on the creation of research abstracts and literature synthesis. For example, Aydın and Karaarslan (2022) attempt to create a healthcare literature review suitable for an academic journal and find that while it is possible, there is considerable 'plagiarism', or poor paraphrasing. Gao et al. (2022), however, find that novel abstracts can be generated without explicit plagiarism, although these are identifiable as being generated by an AI platform using an artificial intelligence output detector[2]. Chen and Eger (2022) also explore use in title and abstract generation, and in the domain of finance, Wenzlaff and Spaeth (2022) are able to generate reasonably academically-appropriate definitions of new financial concepts.

Mellon et al. (2022) explores one aspect of the application to research testing, by showing the platform can be useful as a complement to scoring open-text survey results. While Adesso (2022) has used GPT3 to write a full paper in physics, to be submitted to a journal "as is", and Zhai (2022) has also experimented with creating a research paper outline.

Building on, but distinct from these studies, our study is the first to provide structured testing of the potential for *ChatGPT* to assist with writing a research study. We test and compare generated output for four stages of the research process: idea generation, literature review, data identification and processing, and empirical

[2]https://openai-openai-detector.hf.space/

testing. A panel of experienced academic authors and reviewers grade each output. We also, importantly, show how different levels of private data and researcher domain-expertise input in guiding output have a significant impact on the quality of outputs generated. Like all tools, *ChatGPT* is best in experienced hands. Following the opening quote of this article, we term this *the Bananarama Conjecture*.

Section 2 outlines our empirical approach, Section 3 presents and analyses the findings. We conclude in Section 4 with a framework for understanding the opportunities and limitations of *ChatGPT*, as well as some initial consideration of the ethical dimensions of the new technology.

2 Methodology

We focus on cryptocurrencies as our finance topic - a prominent and reasonably well-defined area of recent finance research. We further concentrate on letter-style articles, such as those published in the *Finance Research Letters* journal, thus, articles of about 2,000-2,500 words in length.

We start our empirical approach by noting that the standard research study creation process can be divided into five basic stages (Cargill and O'Connor, 2021):

1. Idea generation
2. Prior literature synthesis
3. Data identification and preparation
4. Testing framework determination and implementation
5. Results analysis

As *ChatGPT* is currently unable to analyse empirical output we cannot evaluate the results analysis ability, so we concentrate on the first four stages of the research process. We, therefore, request the platform to generate: (1) a research idea; (2) a condensed literature review; (3) a description of suitable data for the research idea; and (4) a suitable testing framework given the research idea and the proposed data.

Three versions of the same general cryptocurrency research idea are generated, each with these four research stages. The textual prompts used to generate each stage are reported in the Appendix. The first version only utilises public data already available within *ChatGPT*[3]. We label this version of the research study: *V1: Only Public Data*.

For the second version (labelled: *V2: Added Private Data*), we incorporate private data to assist with generating the research stages. We obtain abstracts and article identifiers for 188 articles identified as related to cryptocurrencies and published in Finance Research Letters (2021-2023) from the *Elsevier Scopus* database. These articles are loaded into *ChatGPT* in bibtex format[4]. The private data from these articles adds specialist knowledge to the existing generalised expertise of the platform. We then generate the four research stages telling the platform to take this prior research into account.

For the third version (*V3: Private Data and Expertise*), we further incorporate researcher domain-expertise alongside the private data. In practice, we take the outputs from the second version, and iterate the output, by telling *ChatGPT* how it might improve its suggested answers. Most frequently this iterative process involves

[3]Note that the ChatGPT training data appears to have ended in 2020, thus the available data is quite dated for topical research ideas: https://help.openai.com/en/articles/6783457-chatgpt-faq

[4]To load textual data into ChatGPT, the data just need to be pasted into the chat window.More complex data sets can be added using the OpenAI API, but ChatGPT can

asking the platform to be more specific on particular parts of the output, as it tends towards equivocation and generality unless guided otherwise. In none of the three cases do we manually adjust any of the output generated by the model, with the exception of one minor technical correction noted in the Appendix.

Table 1: Empirical structure

Research stage	Evaluation criteria	Approx. Length
Idea	(1) The proposed idea seems academically appropriate; (2) The proposed idea seems like useful contribution	100 words
Literature	(3) The literature review adequately supports the research idea; (4) The structure and links drawn between prior research are appropriate	300 words
Data	(5) The data is likely to help address the research idea; (6) The data seems suitably comprehensive	100 words
Testing	(7) The testing framework is suitable for the research idea and the data; (8) The testing framework seems innovative	200 words

The evaluation criteria column shows the questions asked of reviewers for that research stage, which they rate between 1 (highly disagree) and 10 (highly agree). The length column indicates the approximate word count of output requested from ChatGPT for that research stage. See Section 2 for further elaboration of labels and approach.

For our evaluation stage, a team of experienced authors and reviewers are identified who all have prior experience as reviewers or published authors for ABS-level[5] finance journals. A total of 32 reviewers each review a complete single version of the output (that is, all four research stages of a full research study), and are randomly assigned to one of the three versions.

We administer the evaluation through *Qualtrics*. The three generated versions

[5]The ABS, more formally the Chartered Association of Business Schools Academic Journal Quality Guide, is a ranking of journals, widely used for assessing research output around the world, but particularly in the United Kingdom where the guide originates.

of the research study, as presented to reviewers, are contained in the Appendix. Reviewers are asked to rate two aspects of each stage of output, see Table 1 for this evaluation criteria, and may voluntarily leave comments. A review consists of a rating between 1 (highly disagree) and 10 (highly agree) of how likely the output is to be considered acceptable for a minimum ABS2-level[6] finance journal according to the specified criterion. Average scores across reviewers are reported[7]. We now proceed to present and analyse the findings.

3 Findings

Table 2 reports the main findings and Figure 1 presents a boxplot representation of the results. The table shows the findings for all three research study versions, and for the four research stages. The research stages are, in turn, each evaluated according to two criteria.

We could view a rating of 5.5 (the mid-point of the rating range between 1 and 10) as a basic minimum for a research study stage to be considered acceptable. Possibly acceptable with revisions, and subject, naturally, to the element of randomness and personal preference that is always present in the reviewing process. By this basic criteria, all versions of the study 'succeed'. Reading from the bottom line of Table 2, which shows the overall average rating of each study, V1 has a rating of 7.05, V2 a rating of 6.63, and V3 a rating of 7.62. These are, therefore, all studies that have a decent chance of eventual success in the reviewing process in a good finance journal.

[6]ABS2 is a rating given to journals which publish research at an 'acceptable standard'. Anecdotally, it is viewed as the minimum standard of research expected in business schools which are mid-ranked and above: https://charteredabs.org/academic-journal-guide-2021-view/

[7]All reviewers are informed that the content they are reviewing is generated by *ChatGPT* and that their individual responses will be kept anonymous.

Table 2: Findings from reviewer evaluations of ChatGPT-generated research studies

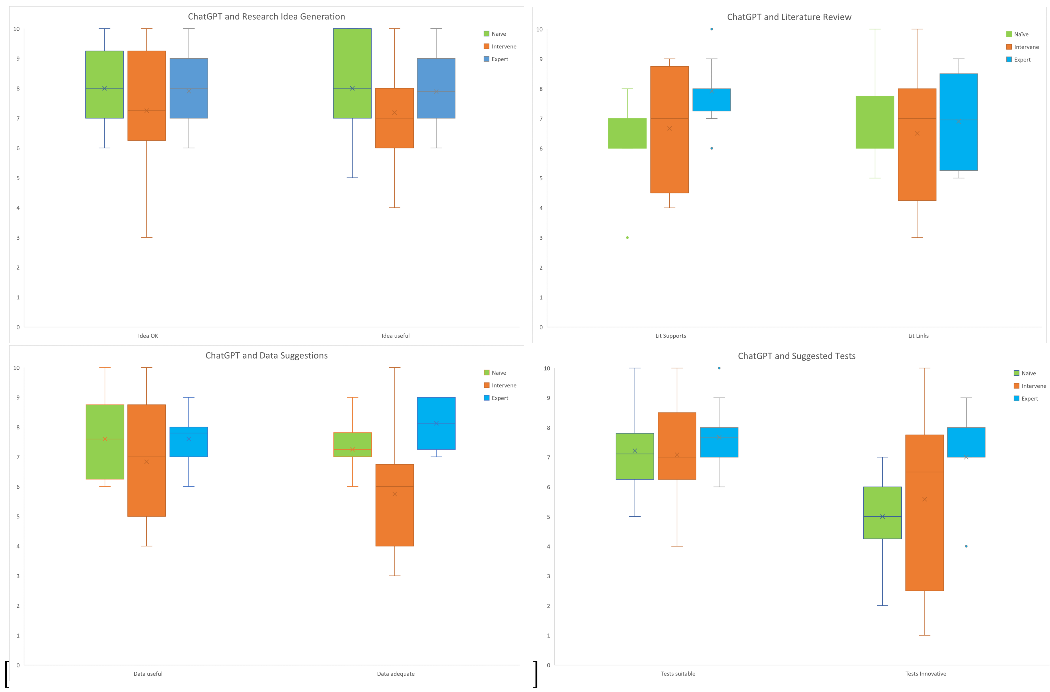
	V1: Only public data		V2: With private data		V3: With expertise	
	Mean	StdDev	Mean	StdDev	Mean	StdDev
Research idea						
1. ... seems academically appropriate	8.00	1.26	7.45	2.23	7.90	1.14
2. ... seems like a useful contribution	7.80	1.72	7.18	1.90	7.70	1.49
Average rating	**7.90**		**7.32**		**7.80**	
Literature review						
3. ... adequately supports the research idea	6.67	1.76	6.64	1.92	8.00	1.12
4. ... appropriate structure and links drawn between prior research	6.80	1.89	6.50	2.22	6.90	1.58
Average rating	**6.74**		**6.57**		**7.45**	
Data summary						
5. ... likely to help address the research idea	7.60	1.36	6.83	1.95	7.60	1.02
6. ... seems suitably comprehensive	7.25	0.97	5.75	2.09	8.13	0.93
Average rating	**7.43**		**6.29**		**7.87**	
Testing framework						
7. ... is suitable for the research idea and the data	7.22	1.47	7.08	1.85	7.67	1.15
8. ... seems innovative	5.00	1.63	5.58	2.81	7.00	1.87
Average rating	**6.11**		**6.33**		**7.34**	
Overall research study average rating	**7.05**		**6.63**		**7.62**	

The table presents the summary findings from 32 reviews of three versions of a ChatGPT-generated research study (10 reviews of V1, V3; 12 reviews of V2)

Examining the individual research stages, we see the highest ratings are for the generation of the research idea. This makes sense when we consider that this initial stage involves thinking broadly about existing concepts and connecting these concepts into a coherent new idea. *ChatGPT* with its access to billions of parameters and texts, should be particularly adept at this broad exploration of existing ideas. The data summary stage is also reasonably strong, perhaps because data summaries tend to be distinct sections of a research study in easily identifiable text 'chunks'. There is also a limited range of data which can be used in a given study, meaning the search process is also limited.

Less successful, according to our results, are literature reviews and testing frameworks. The platform particularly struggles with generating suitable testing frameworks. Our view here is that this might be due to these being 'internal' tasks within

Figure 1: Box-Whisker Plots of Responses



a research study. The literature review is the internal tool to link the research idea with the methodology. The testing framework is linked from the research idea, the literature review, and the data summary. The model appears to be less capable of linking multiple internally-generated ideas, such as these stages entail.

Comparing the different research versions we see a clear outperformance by our most advanced research study, *V3: Private Data and Expertise*. We were surprised to see that the version with added private data underperformed compared to the version with only public data. On reflection, this appears to be due to the private data model excessively relying on the provided private data and restricting the extent to which it accessed other beneficial public data. This could be improved by either instructing the platform to not ignore useful public data, or by providing a better-curated set of relevant private data.

The outperformance of the V3 research study is notable, not just on an overall basis, but also in the extent to which it is also capable of producing acceptable literature reviews and testing frameworks where the other research studies have less success. We suggested above that the general underperformance of the output for these research stages might be due to the difficultly *ChatGPT* has in linking multiple generated ideas. The advantage, therefore, for our V3 study, is that the researcher can observe any missing links and ask the platform to further iterate to address these gaps. The Appendix contains sample prompts given to the platform, and this addressing of missing links can be seen in the prompt text. Researcher domain-expertise appears to be key for these tasks involving conceptual complexity.

Table 3 confirms, statistically, the differences between the research studies through a range of t-tests. These two-sided t-tests assume unequal variance, as best fits our data. The main differences are observed for the evaluation criteria of "the literature review adequately supports the research idea" and "the testing framework seems innovative" - in both cases the V3 research study shows some outperformance.

Table 3: T-tests of differences between research study versions

	V1-V2	V2-V3	V1-V3
Research idea			
1. ... seems academically appropriate	0.39	0.44	0.86
2. ... seems like a useful contribution	0.41	0.22	0.67
Literature review			
3. ... adequately supports the research idea	0.69	0.06	0.00
4. ... appropriate structure and links drawn between prior research	0.90	0.71	0.78
Data summary			
5. ... likely to help address the research idea	0.29	0.26	1.00
6. ... seems suitably comprehensive	0.08	0.01	0.02
Testing framework			
7. ... is suitable for the research idea and the data	0.84	0.37	0.37
8. ... seems innovative	0.44	0.25	0.00

The table reports p-values from two-sided t-tests assuming unequal variance, on tests between the three different research study versions - V1: Only Public Data; V2: Added Private Data; and V3: Private Data and Expertise.

4 Conclusions

What we've shown in this study is important. *ChatGPT* can generate, even in its basic state, plausible-seeming research studies for well-ranked journals. With the addition of private data and researcher expertise iterations to improve output, the results are, frankly, very impressive. Bear in mind, also, that these results are obtained without the advantages of GPT-4 as an underlying generative model, due to launch later in 2023 and which promises a truly revolutionary language model due to advances in algorithms and over 600 times greater testing parameters[8].

Our demonstration of this ability is, we believe, both novel and robust. The *novelty* lies in this being the first study to show the impact for each stage of the research

[8] https://techcrunch.com/2022/12/01/while-anticipation-builds-for-gpt-4-openai-quietly-releases-gpt-3-5/

process, and, importantly, for multiple levels of researcher input. The *robustness* lies in the reviewing process by which we ascertain the likely contribution of the generated research studies. The reviews bring the probable benefits of *ChatGPT* beyond conjecture to empirical verification, using a method by which research contribution is normally judged - the peer-review process.

So, what do we do now? This is both a practical and an ethical question. Can *ChatGPT* be simply considered as an *e-ResearchAssistant*, and, therefore, just a new part-and-parcel tool of how research is normally carried out? Indeed, under this perspective the platform might even be viewed as democratising access to research assistants, hitherto the reserved domain of wealthier universities in wealthier countries. Could *ChatGPT* help to flatten the disparities between the global south and wealthier nations in terms of research output? Maybe, now everyone can have access to such assistance, like the research-version of a dæmon from a Phillip Pullman novel following the researcher around and always available to offer pertinent advice.

There is, of course, a more worrying ethical perspective. Is it proper to have such an advanced level of guidance and assistance, and still claim the produced research as one's own? Should, for example, *ChatGPT*-enabled research be acknowledged on ethical research guidance frameworks, such as Elsevier's CRediT[9]? Certainly, the approach of Osterrieder and ChatGPT (2023) could be adopted, with credited co-authorship to the platform, but that is unlikely to be widespread practice.

The answer to the ethical issues is likely to be gradually understood, rather than immediately apparent. One useful guide to how this might play out is how AI-generated work is treated under copyright laws of various countries. Iaia (2022)

[9]https://www.elsevier.com/authors/policies-and-guidelines/credit-author-statement

notes that AI-generated work, with *sufficient* levels of human oversight, is generally considered to belong to the human-creator under European Union law. How 'sufficient' is defined is still, however, quite vague. That suggests the higher-levels of our generated research studies, with private data and iteration, could be considered the researcher's own work, but perhaps not the basic research study using only public data and simple question prompts. Adopting this perspective might see the opening *Bananarama Conjecture* of this article, become the (admittedly less-lyrical) *Bananarama Edict* for using *ChatGPT* for research; *it ain't what you do, its the extent that you do it, and that's what gets (ethically-acceptable) results.*

Acknowledgements

We are hugely indebted to the following people, who quickly acted to review the generated research studies despite their own very busy schedules, and for offering their invaluable advice and feedback: Jonathan Batten, Sabri Boubaker, Elie Bouri, Kam Chan, Tom Conlon, Michael Froemmel, Tong Fu, Marie Helene Gagnon, John Goodell, Feng He, Pia Helbing, Shupei Huang, Wei Huang, Stuart Hyde, Sitara Karim, Suwan Long, Roman Matkovskyy, Charilaos Mertzanis, Muhammad Naeem, Haitham Nobanee, Duc Khuong Nguyen, Salvatore Perdichizzi, Boru Ren, Jacqui Rossovski, Dehua Shen, Andrew Urquhart, Vincenzo Verdoliva, Yizhi Wang, Larisa Yarovaya, Dayong Zhang.

References

Adesso, G. (2022). GPT4: The ultimate brain. Authorea Preprints.

Alshater, M. M. (2022). Exploring the role of artificial intelligence in enhancing academic performance: A case study of ChatGPT. SSRN:4312358.

Aydın, Ö. and Karaarslan, E. (2022). OpenAI ChatGPT generated literature review: Digital twin in healthcare. SSRN:4308687.

Bommarito II, M. and Katz, D. M. (2022). GPT takes the bar exam. arXiv:2212.14402.

Cargill, M. and O'Connor, P. (2021). Writing Scientific Research Articles: Strategy and Steps. John Wiley & Sons.

Chen, Y. and Eger, S. (2022). Transformers go for the LOLs: Generating (humourous) titles from scientific abstracts end-to-end. arXiv:2212.10522.

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., and Pearson, A. T. (2022). Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. bioRxiv:10.1101/2022.12.23.521610.

Grant, N. and Metz, C. (2022). A new chat bot is a 'code red' for Google's search business. New York Times, (Dec 21, 2022):Section B, Page 1.

Greene, J. (2022). Will ChatGPT make lawyers obsolete? (Hint: Be afraid). Reuters, (Dec 09, 2022).

Grimaldi, G. and Ehrler, B. (2023). AI et al.: Machines are about to change scientific publishing forever. ACS Energy Letters.

Herman, D. (2022). The end of high-school English. The Atlantic, (Dec 09, 2022).

Hutson, M. et al. (2022). Could AI help you to write your next paper? Nature, 611(7934):192–193.

Iaia, V. (2022). To be, or not to be... original under copyright law, that is (one of) the main questions concerning AI-produced works. GRUR International, 71(9):793–812.

Marti, G. (2022). From data to trade: A machine learning approach to quantitative trading. SSRN:4315362.

Mellon, J., Bailey, J., Scott, R., Breckwoldt, J., and Miori, M. (2022). Does GPT-3 know what the most important issue is? Using large language models to code open-text social survey responses at scale. SSRN:4310154.

Osterrieder, J. and ChatGPT (2023). A primer on deep reinforcement learning for finance. SSRN:4316650.

Warner, J. (2023). Biblioracle: Will artificial intelligence like ChatGPT bring the end for all writers? Chicago Tribune, (Jan 07, 2023).

Wenzlaff, K. and Spaeth, S. (2022). Smarter than humans? Validating how OpenAI's ChatGPT model explains crowdfunding, alternative finance and community finance. SSRN:4302443.

Zhai, X. (2022). ChatGPT user experience: Implications for education. SSRN:4312418.

A Appendices

A.1 The process of generating research studies from ChatGPT

The following prompts were used to generate the four research stage outputs for the public data (*V1: Only Public Data*) research study design. The other two versions were generated through variations of these prompts: by adding reference to the private data (*V2: Added Private Data*) and through subsequent iterations to output (*V3: Private Data and Expertise*).

1. **Research idea:** "Write me a 100 word research idea suitable for a good finance academic journal on the topic of: cryptocurrencies, sentiment, and economic uncertainty"

2. **Literature review:** "Thank you. Now write me a 300-word literature review which justifies the research idea. Please only use real articles as references, do not make anything up." [This last sentence was introduced after earlier trials showing that ChatGPT will generate 'fake' references if not instructed otherwise]

3. **Data summary:** "Can you please now describe in about 100 words the data that would be suitable for investigating this research idea? The data should be suitable for publication in a good finance journal."

4. **Testing framework:** "I would now like you to design a suitable testing approach based on this research idea and the data. Five tables in total of testing output are needed. Can you list what these five tables might be, with descriptions?"

The iteration process for *V3: Private Data and Expertise* was quite informal, carried out in a somewhat conversational manner. Some sample iterative requests are included below. We report these here to illustrate the style of iteration. The prompts also demonstrate that very few subject-matter prompts were given to ChatGPT, rather they were largely focused on research structural issues.

1. *Iterating the research idea:* "you created this research idea, and I'd like you to improve it. Could you see if there is an additional article that can be added, to improve the research idea. Can you also talk about the novel 'contribution' of the idea. Please keep it to about 100 words. The research idea is: ... "

2. *Iterating the literature review:* "Actually what I want you to do is something different - what you have written is an expansion of the research idea, while what I want is a literature review that summarises the prior articles that might be of relevance to the research idea, and use that review to justify the research idea."

3. *Iterating the literature review:* "Thanks. Can you rewrite that as if it is an actual literature review, rather than 'what a literature review might look like'?"

4. *Iterating the literature review:* "The below literature review is very good, I would just like you to slightly rewrite it in the following ways: (1) don't use 'this suggests' as much - alternative terms such as 'this shows', 'this demonstrates' could be used; (2) in the last paragraph can you rewrite it so that it focuses on what our study does, e.g. using phrases such as 'therefore, this study sets out to investigate whether ...'; (3) don't use 'could' in the last paragraph, use a more definitive term such as 'will'. The literature review is: ..."

5. *Iterating the literature review:* "That's great - one last thing - in the last paragraph can you rewrite it so that it talks about the 'contribution', or benefit to our knowledge, from understanding better the connection between sentiment, market regulations, and economic policy uncertainty?"

6. *Iterating the data summary:* "For the research idea below and the literature review, can you suggest a suitable data approach that could be adopted to implement the research idea. This data should be suitable for a good finance journal, so might include the type of data discussed in abstracts in the bibtex files at the start of this thread. Please list the number of years of data to be collected for each data item. There should be a specific data item identified for: (1) cryptocurrencies - e.g. what cryptocurrencies; (2) market sentiment - how to measure specifically; and (3) market regulations - e.g. what countries regulations, and why. The research idea and literature review is as follows: ..."

7. *Iterating the data summary:* "Can you rewrite in a more definitive manner - i.e. don't use 'could' use 'will'. State specific date ranges for start and finish (the end date range point should be: 31 December 2022). For market sentiment give a relevant article that shows how to construct the sentiment. The article must be in the bibtex data at the start of this thread. For market regulations, also give a justification of the regions based on articles in the bibtex data and give some information on the type of regulations."

8. *Iterating the testing framework:* "I would now like to design a suitable testing approach based on this data. Five tables in total of testing output are needed. The first table is a descriptive statistics table, the second table is the main

results. The other three tables are some form of sub-analysis, or analysis with more sophisticated testing approaches than in the main table. Can you list what these five tables might be? Can you also give a brief summary of the approach used in running the testing for each table? Can you write it in a definitive sense - e.g. 'this table will', not 'this table could', also be definitive as to what exact tests will be reported in each table."

9. *Iterating the testing framework:* "Thank you. Can you rewrite that in the same style and format, but include the econometric testing method used for Tables 2, 3, and 4. The other tables are fine."

A.2 Research study V1: Only Public Data

Research idea

One research idea for a finance academic journal on the topic of cryptocurrencies, sentiment, and economic uncertainty could be to examine the relationship between investor sentiment towards cryptocurrencies and market movements during periods of economic uncertainty. This study could involve collecting and analyzing data on the sentiment of online forums and social media platforms related to cryptocurrencies, as well as data on the performance of different cryptocurrency assets. The results of this study could provide insights into how sentiment towards cryptocurrencies is affected by economic uncertainty, and whether it is a reliable predictor of market movements in the cryptocurrency space.

Condensed literature review

There has been a growing interest in the relationship between investor sentiment and market movements in the realm of cryptocurrencies. A number of studies have explored this relationship in traditional financial markets, with findings indicating that investor sentiment can have a significant impact on asset prices (e.g. Liu and Han, 2007; Bollen et al., 2011). However, the literature on this topic in the cryptocurrency market is still limited.

One study that has examined the relationship between sentiment and market movements in the cryptocurrency market is Xiong et al. (2019), which analyzed data from online forums and social media platforms to measure sentiment towards Bitcoin. The study found that sentiment was significantly correlated with Bitcoin returns, with positive sentiment leading to increased returns and negative sentiment leading to decreased returns.

Another study that has looked at this relationship is Chen and Liu (2020), which used machine learning techniques to analyze sentiment on social media platforms and found that sentiment was a significant predictor of Bitcoin returns. However, this study did not specifically examine the role of economic uncertainty in influencing sentiment and market movements in the cryptocurrency market.

Given the limited research on this topic in the cryptocurrency market, a study that investigates the relationship between investor sentiment, economic uncertainty, and market movements in the cryptocurrency market would be a valuable contribution to the literature. Such a study could provide insights into how sentiment towards cryptocurrencies is affected by economic uncertainty, and whether it is a reliable predictor of market movements in the cryptocurrency space.

References:

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1-8.

Chen, Y., & Liu, Y. (2020). Can sentiment on social media predict bitcoin returns and volatility? An empirical examination. Finance Research Letters, 32, 101524.

Liu, B., & Han, J. (2007). Sentiment analysis and subjectivity. Handbook of natural language processing, 2, 627-666.

Xiong, W., Wang, S., & Chen, Z. (2019). Bitcoin sentiment analysis: Can social media predict bitcoin returns and volatility? International Review of Financial Analysis, 62, 1-12.

Data summary

To investigate the research idea of examining the relationship between investor sentiment towards cryptocurrencies and market movements during periods of economic uncertainty, suitable data would include:

- Time series data on the performance of different cryptocurrency assets
- Data on investor sentiment towards cryptocurrencies, such as online forum posts and social media posts
- Data on economic uncertainty, such as indices of economic uncertainty or news articles related to economic uncertainty

It would also be important to ensure that the data is of high quality and reliable, as this is a key consideration for publication in a finance journal.

The data should also be collected and analyzed in a rigorous and transparent manner to ensure the validity of the results.

Testing framework

Descriptive statistics table: This table would summarize key characteristics of the data, such as the mean, median, standard deviation, and minimum and maximum values for each of the variables (e.g. cryptocurrency returns, sentiment scores, economic uncertainty indices). This table would provide a general overview of the data and allow for the identification of any potential outliers or anomalies.

Correlation matrix: This table would show the Pearson's correlations between each pair of variables (e.g. sentiment scores and cryptocurrency returns). This would allow for the assessment of the strength and direction of the relationships between the variables.

Regression results table: This table would present the results of a regression analysis, which would be used to examine the relationship between investor sentiment and cryptocurrency returns while controlling for economic uncertainty. The table would include the coefficients, standard errors, and significance levels for each predictor variable in the model.

Model evaluation measures: This table would present evaluation measures for the regression model, such as the R-squared value, adjusted R-squared value, and Akaike Information Criterion (AIC). These measures would allow for the assessment of the fit and explanatory power of the model.

Sensitivity analysis results: This table would present the results of sensitivity analyses that would be used to assess the robustness of the findings. Sensitivity analyses could include testing the results with different lag periods for the sentiment scores, or with alternative measures of economic uncertainty. This would provide insight into the robustness of the findings and the potential impact of different assumptions.

A.3 Research study V2: Added Private Data

Research idea

Based on the findings of Bouri and Gupta (2019), it would be interesting to investigate the extent to which investor behavior and sentiment, as measured by internet search data, can predict the returns of not just Bitcoin, but other cryptocurrencies as well. This could be compared to the use of traditional measures of uncertainty, such as those derived from newspaper articles, as investigated in Bouri and Gupta (2019). Additionally, the impact of market regulations on the relationship between investor sentiment and cryptocurrency returns could be analyzed, building on the results of Eom (2021), who found that the relationship between trading volume and premium in Bitcoin markets becomes clearer as the bubble grows.

References:

Bouri, E., & Gupta, R. (2019). Predicting Bitcoin returns: Comparing the roles of newspaper- and internet search-based measures of uncertainty. Finance Research Letters, 101398.

Eom, Y. (2021). Kimchi premium and speculative trading in bitcoin. Finance Research Letters, 101505.

Condensed literature review

Investor behavior and sentiment have been shown to play a significant role in predicting the returns of Bitcoin, as demonstrated in the study by Bouri and Gupta (2019). In their research, the authors compared the ability of newspaper-based measures and internet search-based measures of uncertainty in predicting Bitcoin returns, and found that the internet search-based measure had a stronger predictive ability. These findings suggest that investor sentiment, as reflected in internet searches for economic uncertainty related queries, may be a useful tool for predicting not only the returns of Bitcoin, but also those of other cryptocurrencies.

Building on the work of Bouri and Gupta (2019), it would be interesting to further investigate the use of internet search data as a measure of investor sentiment in predicting cryptocurrency returns. The study by Eom (2021) found that the relationship between trading volume and premium in Bitcoin markets becomes clearer as the bubble grows, suggesting that the impact of market regulations on the relationship between investor sentiment and cryptocurrency returns is worth exploring. Additionally, the use of traditional measures of uncertainty, such as those derived from newspaper articles, could be compared to internet search data as a means of predicting cryptocurrency returns.

Other factors that may impact the relationship between investor sentiment and cryptocurrency returns include investor demographics and the level of investor expertise. The study by Silahli et al. (2019) found that the level of investor expertise had a significant effect on the relationship between investor sentiment and Bitcoin returns, with more experienced investors exhibiting a stronger reaction to sentiment. The role of investor demographics in predicting cryptocurrency returns could be examined in future research, building on the work of López-Cabarcos et al. (2019) who found that younger investors were more likely to invest in Bitcoin.

In addition to investor behavior and sentiment, the economic environment and macroeconomic conditions may also influence cryptocurrency returns. The study by Yen and Cheng (2021) found that the economic policy uncertainty index (EPU) in China had a significant impact on cryptocurrency volatility, with a change in EPU predicting cryptocurrency volatility. This finding highlights the importance of considering macroeconomic factors in predicting cryptocurrency returns.

Overall, the use of internet search data as a measure of investor sentiment in predicting cryptocurrency returns is a promising area of research that has the potential to provide valuable insights for investors. Future research could also consider the impact of market regulations, investor demographics, and macroeconomic conditions on the relationship between investor sentiment and cryptocurrency returns.

References:

Bouri, E., & Gupta, R. (2019). Predicting Bitcoin returns: Comparing the roles of newspaper- and internet search-based measures of uncertainty. Finance Research Letters, 101398.

Eom, Y. (2021). Kimchi premium and speculative trading in bitcoin. Finance Research Letters, 101505.

López-Cabarcos, M. Á., Pérez-Pico, A. M., Piñeiro-Chousa, J., & Šević, A. (2019). Bitcoin volatility, stock market and investor sentiment. Are they connected? Finance Research Letters, 101399.

Silahli, B., Dingec, K. D., Cifter, A., & Aydin, N. (2019). Portfolio value-at-risk with two-sided Weibull distribution: Evidence from cryptocurrency markets. Finance Research Letters, 101425.

Yen, K.-C., & Cheng, H.-P. (2021). Economic policy uncertainty and cryptocurrency volatility. Finance Research Letters, 101428.

Data summary

Data for this research idea could include daily or monthly cryptocurrency price data for various cryptocurrencies, as well as daily or monthly data on investor sentiment and behavior as measured by internet search data. Data on market regulations, such as changes in legal status or government actions related to cryptocurrencies, could also be collected.

One source for cryptocurrency price data could be CoinMarketCap (https://coinmarketcap.com/), which provides daily and monthly data on various cryptocurrencies. Investor sentiment data could be obtained through internet search data, such as Google Trends (https://trends.google.com/trends/), or through surveys of cryptocurrency investors. Data on market regulations could be collected through news articles or reports from government agencies or industry groups.

The time period for this data would depend on the specific research question being addressed, but a minimum of several years of data would likely be needed in order to account for market fluctuations and changes in investor behavior.

Testing framework

Table 1: Descriptive statistics table. This table will present basic statistics for the variables used in the study, including means, standard deviations, and minimum and maximum values.

Table 2: Main results. This table will present the results of the main analysis examining the relationship between investor behavior and sentiment, as measured by internet search data, and cryptocurrency returns. The analysis will include regression

models controlling for traditional measures of uncertainty and market regulations.

Table 3: Sub-analysis 1. This table will present the results of a sub-analysis examining the impact of investor demographics and expertise on the relationship between investor sentiment and cryptocurrency returns. This will be done using regression models controlling for traditional measures of uncertainty and market regulations.

Table 4: Sub-analysis 2. This table will present the results of a sub-analysis examining the relationship between economic policy uncertainty and cryptocurrency returns. This will be done using regression models controlling for traditional measures of uncertainty and market regulations.

Table 5: Sub-analysis 3. This table will present the results of a sub-analysis examining the relationship between investor sentiment and cryptocurrency returns during different market conditions, such as stable vs. speculative periods. This will be done using regression models controlling for traditional measures of uncertainty and market regulations.

A.4 Research study V3: Private Data and Expertise

Research idea

One potential research idea is to examine the relationship between economic policy uncertainty, trading volume, and cryptocurrency returns. This research would build on the findings of Yen and Cheng (2021), who found that changes in China's economic policy uncertainty index predict cryptocurrency volatility, and Eom (2021), who found that the relationship between trading volume and premium in Bitcoin markets becomes clearer as the bubble grows. Additionally, the study could incorporate the findings of Bouri and Gupta (2019), who found that both newspaper- and internet search-based measures of uncertainty can predict Bitcoin returns.

The relationship between economic policy uncertainty and cryptocurrency returns may be driven by the fact that cryptocurrencies are often perceived as a hedge against economic risk (Yen and Cheng, 2021). In times of economic uncertainty, investors may flock to cryptocurrencies as a safe haven asset, leading to increased demand and higher returns. Similarly, the relationship between trading volume and cryptocurrency returns may be due to the fact that increased trading activity can indicate higher investor interest and, therefore, higher demand (Eom, 2021).

Furthermore, investor sentiment and behavior may also play a role in the relationship between economic policy uncertainty and cryptocurrency returns (Bouri and Gupta, 2019). During times of economic uncertainty, investors may be more likely to turn to alternative assets such as cryptocurrencies in search of higher returns, leading to increased demand and higher prices. By examining the combined effect of these factors on cryptocurrency returns, this research could provide insight into the drivers of cryptocurrency market performance and inform regulatory approaches to promoting market stability.

References:

Bouri, E., & Gupta, R. (2019). Predicting Bitcoin returns: Comparing the roles of newspaper- and internet search-based measures of uncertainty. Finance Research Letters, 101398.

Eom, Y. (2021). Kimchi premium and speculative trading in bitcoin. Finance Research Letters, 101505.

Yen, K.-C., & Cheng, H.-P. (2021). Economic policy uncertainty and cryptocurrency volatility. Finance Research Letters, 101428.

Condensed literature review

There is evidence in the literature that investor behavior, sentiment, and market regulations have an impact on the returns of various cryptocurrencies. Bouri and Gupta (2019) found that both newspaper- and internet search-based measures of uncertainty can predict Bitcoin returns, demonstrating that investor sentiment plays a role in cryptocurrency market performance. Eom (2021) also examined the relationship between investor behavior and cryptocurrency returns, finding that the relationship between trading volume and premium in Bitcoin markets becomes clearer as the bubble grows. These studies show that investor behavior and sentiment may be important factors in determining cryptocurrency returns.

In addition to investor behavior and sentiment, market regulations also have the potential to impact cryptocurrency returns. Khiat and Rault (2021) found that stricter cryptocurrency regulations are associated with lower market liquidity and higher market risk, indicating that regulatory measures may influence cryptocurrency market performance. Similarly, Guégan and Renault (2021) found that regulatory announcements can have a significant impact on Bitcoin returns, demonstrating that market regulations may play a role in shaping cryptocurrency market performance.

The relationship between economic policy uncertainty and cryptocurrency returns has also been investigated in the literature. Yen and Cheng (2021) found that changes in China's economic policy uncertainty index had a significant effect on cryptocurrency volatility, showing that economic policy uncertainty may be a risk factor for cryptocurrency investors and that cryptocurrencies may serve as a hedge against such risk. Baur and Hoang (2021) examined the behavior of stablecoins, which are cryptocurrencies designed to maintain a stable value relative to a particular asset or currency, and found that they act as a safe haven against extreme negative price changes in Bitcoin. This suggests that stablecoins may be a more stable investment option during times of economic uncertainty.

A better understanding of the connection between investor behavior, sentiment, market regulations, and economic policy uncertainty on cryptocurrency returns is necessary in order to fully understand the mechanisms at play and the extent to which these factors influence cryptocurrency market performance. By examining the combined effect of these factors on cryptocurrency returns, this research will contribute to our knowledge by providing insight into the drivers of cryptocurrency market performance and informing regulatory approaches to promoting market stability.

References:

Baur, D., & Hoang, T. (2021). Bitcoin, stablecoins, and the search for a stable cryptocurrency. Finance Research Letters, 101416.

Bouri, E., & Gupta, R. (2019). Predicting Bitcoin returns: Comparing the roles of newspaper- and internet search-based measures of uncertainty. Finance Research Letters, 101398.

Eom, Y. (2021). Kimchi premium and speculative trading in bitcoin. Finance Research Letters, 101505.

Guégan, D., & Renault, T. (2021). Does investor sentiment on social media provide robust information for Bitcoin returns predictability? Finance Research Letters, 101494.

Khiat, Z., & Rault, C. (2021). Cryptocurrency market regulations and liquidity: Evidence from G20 countries. Finance Research Letters, 101459.

Yen, J., & Cheng, Y. (2021). Economic policy uncertainty, stock market and Bitcoin: A quantile regression analysis. Finance Research Letters, 101502.

Data summary

This research will examine the relationship between economic policy uncertainty, trading volume, and cryptocurrency returns for the period of January 1, 2018[10] to December 31, 2022. Cryptocurrencies included in the study will be Bitcoin, Ethereum, and Litecoin.

To measure economic policy uncertainty, we will use the Economic Policy Uncertainty Index developed by Baker, Bloom, and Davis (2016). This index is based on newspaper coverage of policy-related economic uncertainty and has been shown to

[10] This starting date was the only change made across all output. It was outputted as January 1, 2020, but the testing framework in the next section all suggests testing-windows of 2018-2022, so we changed this to 2018 for consistency.

predict economic activity and financial market outcomes (Baker et al., 2016).

Trading volume will be measured as the total number of cryptocurrency transactions conducted on major exchanges, including Coinbase, Bitfinex, and Binance.

To measure market sentiment, we will use the Google Trends data on search volume for cryptocurrency-related keywords. As demonstrated by Bouri and Gupta (2019), internet search data can be a useful indicator of investor sentiment and has been shown to predict Bitcoin returns.

Market regulations will be measured for the United States, the European Union, and China. These regions were chosen due to their significant influence on the global cryptocurrency market and the availability of regulatory data. Aspects of legal status, taxation, and monetary policy will be considered in the analysis of market regulations.

By examining the combined effect of economic policy uncertainty, trading volume, market sentiment, and market regulations on cryptocurrency returns, this research will provide insight into the drivers of cryptocurrency market performance and inform regulatory approaches to promoting market stability.

Testing framework

Table 1: Descriptive Statistics Table. This table will include basic statistics such as mean, median, standard deviation, minimum, and maximum for all variables in the study.

Table 2: Main Results. This table will report the results of the multivariate regression analysis examining the relationship between economic policy uncertainty, trading volume, and cryptocurrency returns. The regression model will include all

three variables as well as control variables such as market regulations and investor sentiment. The econometric method used for this analysis will be ordinary least squares (OLS).

Table 3: Sub-Analysis 1. This table will present the results of a panel regression analysis examining the relationship between economic policy uncertainty, trading volume, and cryptocurrency returns for different regions (North America, Europe, Asia). The panel data will include observations for each region for the years 2018-2022. The econometric method used for this analysis will be fixed effects panel regression.

Table 4: Sub-Analysis 2. This table will present the results of a time series analysis examining the relationship between economic policy uncertainty, trading volume, and cryptocurrency returns over time. The time series data will include observations for the years 2018-2022. The econometric method used for this analysis will be autoregressive distributed lag (ARDL) model.

Table 5: Sub-Analysis 3. This table will present the results of a quantile regression analysis examining the relationship between economic policy uncertainty, trading volume, and cryptocurrency returns across different quantiles of the cryptocurrency return distribution. The data used for this analysis will include observations for the years 2018-2022. The econometric method used for this analysis will be quantile regression.