

Fieberg, Christian; Hornuf, Lars; Streich, David J.

Working Paper

Using GPT-4 for Financial Advice

CESifo Working Paper, No. 10529

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Fieberg, Christian; Hornuf, Lars; Streich, David J. (2023) : Using GPT-4 for Financial Advice, CESifo Working Paper, No. 10529, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/279279>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Using GPT-4 for Financial Advice

Christian Fieberg, Lars Hornuf, David J. Streich

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: <https://www.cesifo.org/en/wp>

Using GPT-4 for Financial Advice

Abstract

We show that the recently released text-based artificial intelligence tool GPT-4 can provide suitable financial advice. The tool suggests specific investment portfolios that reflect an investor's individual circumstances such as risk tolerance, risk capacity, and sustainability preference. Notably, while the suggested portfolios display home bias and are rather insensitive to the investment horizon, historical risk-adjusted performance is on par with a professionally managed benchmark portfolio. Given the current inability of GPT-4 to provide full-service financial advice, it may be used by financial advisors as a back-office tool for portfolio recommendation.

JEL-Codes: G000, G110.

Keywords: GPT-4, ChatGPT, financial advice, artificial intelligence, portfolio management.

Christian Fieberg
City University of Applied Sciences
Bremen / Germany
christian.fieberg@hs-bremen.de

Lars Hornuf
Dresden University of Technology
Dresden / Germany
lars.hornuf@tu-dresden.de

David J. Streich
KU University Eichstaett-Ingolstadt
Ingolstadt / Germany
david.streich@ku.de

June 23, 2023

Co-authors are listed in alphabetical order.

1 Introduction

Text-based artificial intelligence (AI) tools such as GPT-4 enable tech-savvy laymen to conduct tasks in complex domains without much prior experience. For example, early users marvel at the chatbot’s ability to draw on the vast collection of texts it was trained on to draft high-quality term papers and poems or spot mistakes in computer code ([The New York Times, 2022](#)). [Dowling and Lucey \(2023\)](#) assess the use of GPT-4 in producing finance research and find that it generates plausible research ideas, literature reviews, and test suggestions. Given the remarkable quality of output even in complex and creative problems such as academic research, it stands to reason that service providers relying to a large extent on freely available information will be rendered obsolete by advanced AI tools. In this article, we examine the opportunities that AI tools offer for financial advice providers by systematically investigating the suitability of the financial advice provided by GPT-4.¹

The financial advice domain provides an interesting case study for three reasons. First, financial decision-making is a highly consequential domain as individuals in aging developed economies will have to rely in large parts on the proceeds of their financial savings to fund retirement expenses. The fact that retail investors consistently make costly mistakes when investing their savings makes financial advice a potentially welfare-enhancing endeavor ([Calvet et al., 2007](#)), especially if it is enhanced by AI. Second, textbook modern portfolio theory suggests that uninformed investors are best off following a simple, passive investment strategy that involves buying and holding a broadly diversified market portfolio that reflects their individual risk preferences ([Lintner, 1965](#)). Information on both the implications of modern portfolio theory and available investment products to implement such a portfolio are readily available online. Thus, GPT-4, whose responses are informed by billions of texts from the internet, is well-equipped to dish out financial advice in the modern portfolio theory tradition. Third, publicly available information is the key driver of predictable asset price movements, which implies that AI applications such as GPT-4 may be particularly well-equipped to adjust

¹ We use GPT-4 instead of the probably more notorious ChatGPT, because GPT-4’s deep learning approach leverages more data and more computation to create increasingly sophisticated responses and is already in use by the financial industry. For example, Morgan Stanley wealth management deploys GPT-4 to organize its knowledge base (see <https://openai.com/product/gpt-4>, last accessed May 25, 2023).

portfolio allocation advice to the arrival of new information. While the current version of GPT-4 is trained on data up to September 2021, it is likely that future applications will be able to leverage real-time information in their decision-making. Fourth, the financial advice industry has moved to automated, digital offerings for the mass market (Jung et al., 2018; Rühr et al., 2019), making generative AI solutions in financial advice a natural progression.

2 Method

Regulatory guidelines on the provision of financial advice as outlined in MiFID II stipulate that financial advisors must take into account their client’s individual circumstances when providing financial recommendations. Thus, we define hypothetical investor profiles to assess whether GPT-4 customizes its recommendations to the individual investor. According to modern portfolio theory, an individual’s optimal portfolio is determined solely by her risk preferences.² Subsequent research has distinguished between an investor’s level of *risk capacity*, which is related to future financial obligations that impose a constraint on the investor’s liquidity, and *risk tolerance*, which reflects an individual’s inherent willingness to take risk and is not limited to the financial domain (Frey et al., 2017; ESMA, 2018). We capture differences in risk capacity by varying our investor profiles’ age and investment horizon. Typically, investors will shift portfolios toward less volatile securities the shorter the remaining investment horizon to account for the liquidity constraint imposed by the nearing end of the investment period. We capture differences in risk tolerance by varying our investor profiles’ subjective risk tolerance. Table 1 contains 4 profiles representing different combinations of investment horizon and risk tolerance.³

² Tobin’s two-fund separation theorem, which is still the foundation of portfolio allocation theory and the core of undergraduate finance courses, states that all investors hold some combination of the risk-free asset and a risky market portfolio, which is defined as the most efficient combination of all available securities (Merton, 1969; Tobin, 1958). The exact proportions of the two components are governed by an individual’s risk preferences, with more risk-averse investors holding higher proportions of the risk-free asset.

³ We collected recommendations for a total of 48 investor profiles (see Table 5), representing varying investment horizons, origin countries, and ESG preferences. Due to limited data history in particular for ESG-themed investment products and to ensure a tractable dataset for analysis, we restrict the sample to U.S. investor profiles and vary investment horizon and risk tolerance according to Table 1.

Table 1: Investor profiles

Profile	Age	Investment horizon	Risk tolerance
1	30	40	High
2	30	40	Low
3	60	5	High
4	60	5	Low

Next, we request suitable portfolio allocations from GPT-4 for each of the investor profiles. Since the algorithm is not designed to provide individual investment advice upon request likely for legal reasons, requests need to be formulated in a hypothetical scenario to obtain specific portfolio suggestions. We use the following request text:

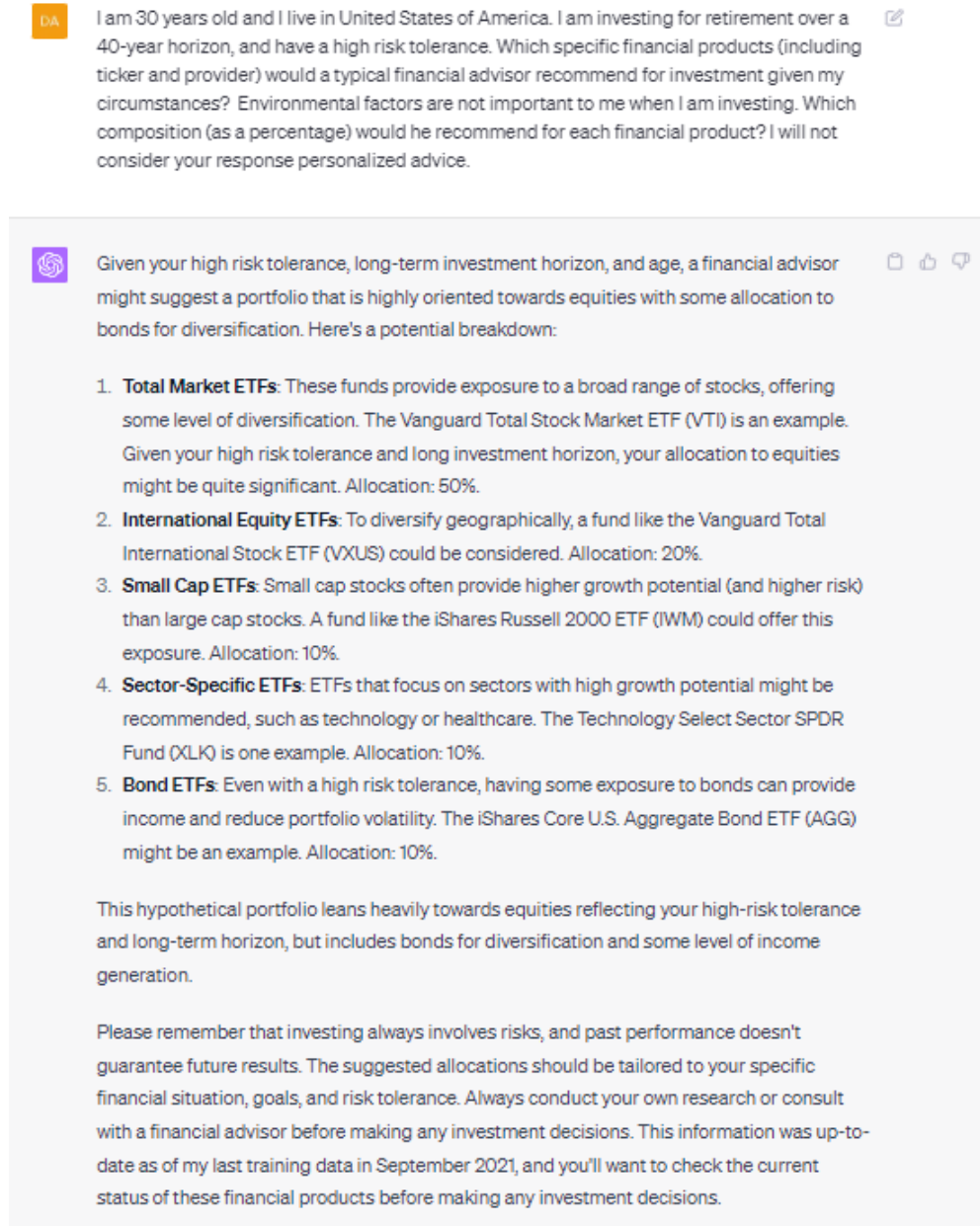
I am [age] years old and live in the United States of America. I am investing for retirement over a [investment horizon]-year horizon, and have a [risk tolerance] risk tolerance. Which specific financial products (including ticker and provider) would a typical financial advisor recommend for investment given my circumstances? Which composition (as a percentage) would he recommend for each financial product? I will not consider your response personalized advice.⁴

The placeholders are replaced with the respective profile’s characteristics as depicted in Table 1. In response to the request, GPT-4 provides a portfolio recommendation consisting of specific investment products and corresponding portfolio shares. As an illustration, Figure 1 displays GPT-4’s response to our request for profile 1. There is a lot to like about the response. First, GPT-4 acknowledges the investor profile’s risk tolerance, investment horizon, and age in determining a suitable portfolio. Second, it suitably recommends a high concentration in equities for this specific profile given the high risk tolerance and long investment horizon. Third, the specific products suggested are all low-cost exchange-traded funds (ETFs) managed by well-known asset managers (Vanguard, BlackRock’s ETF platform iShares, State Street’s

⁴ Since we included information on ESG preferences in our initial investor profiles, requests additionally contained the sentence “Environmental aspects are [not] important to me when I am investing.” All four profiles used in our main analysis had no sustainability preferences.

SPDR platform). Fourth, GPT-4 seems to make an effort to explain the reasoning behind its recommendation, which likely helps improve acceptance (Litterscheidt and Streich, 2020).

Figure 1: GPT-4 portfolio suggestion



The image shows a chat interface with a user message and a GPT-4 response. The user message is in a light blue bubble, and the GPT-4 response is in a white bubble with a purple icon. The response includes a list of five ETF categories with their allocations and a disclaimer at the bottom.

User Query: I am 30 years old and I live in United States of America. I am investing for retirement over a 40-year horizon, and have a high risk tolerance. Which specific financial products (including ticker and provider) would a typical financial advisor recommend for investment given my circumstances? Environmental factors are not important to me when I am investing. Which composition (as a percentage) would he recommend for each financial product? I will not consider your response personalized advice.

GPT-4 Response: Given your high risk tolerance, long-term investment horizon, and age, a financial advisor might suggest a portfolio that is highly oriented towards equities with some allocation to bonds for diversification. Here's a potential breakdown:

- Total Market ETFs:** These funds provide exposure to a broad range of stocks, offering some level of diversification. The Vanguard Total Stock Market ETF (VTI) is an example. Given your high risk tolerance and long investment horizon, your allocation to equities might be quite significant. Allocation: 50%.
- International Equity ETFs:** To diversify geographically, a fund like the Vanguard Total International Stock ETF (VXUS) could be considered. Allocation: 20%.
- Small Cap ETFs:** Small cap stocks often provide higher growth potential (and higher risk) than large cap stocks. A fund like the iShares Russell 2000 ETF (IWM) could offer this exposure. Allocation: 10%.
- Sector-Specific ETFs:** ETFs that focus on sectors with high growth potential might be recommended, such as technology or healthcare. The Technology Select Sector SPDR Fund (XLK) is one example. Allocation: 10%.
- Bond ETFs:** Even with a high risk tolerance, having some exposure to bonds can provide income and reduce portfolio volatility. The iShares Core U.S. Aggregate Bond ETF (AGG) might be an example. Allocation: 10%.

This hypothetical portfolio leans heavily towards equities reflecting your high-risk tolerance and long-term horizon, but includes bonds for diversification and some level of income generation.

Please remember that investing always involves risks, and past performance doesn't guarantee future results. The suggested allocations should be tailored to your specific financial situation, goals, and risk tolerance. Always conduct your own research or consult with a financial advisor before making any investment decisions. This information was up-to-date as of my last training data in September 2021, and you'll want to check the current status of these financial products before making any investment decisions.

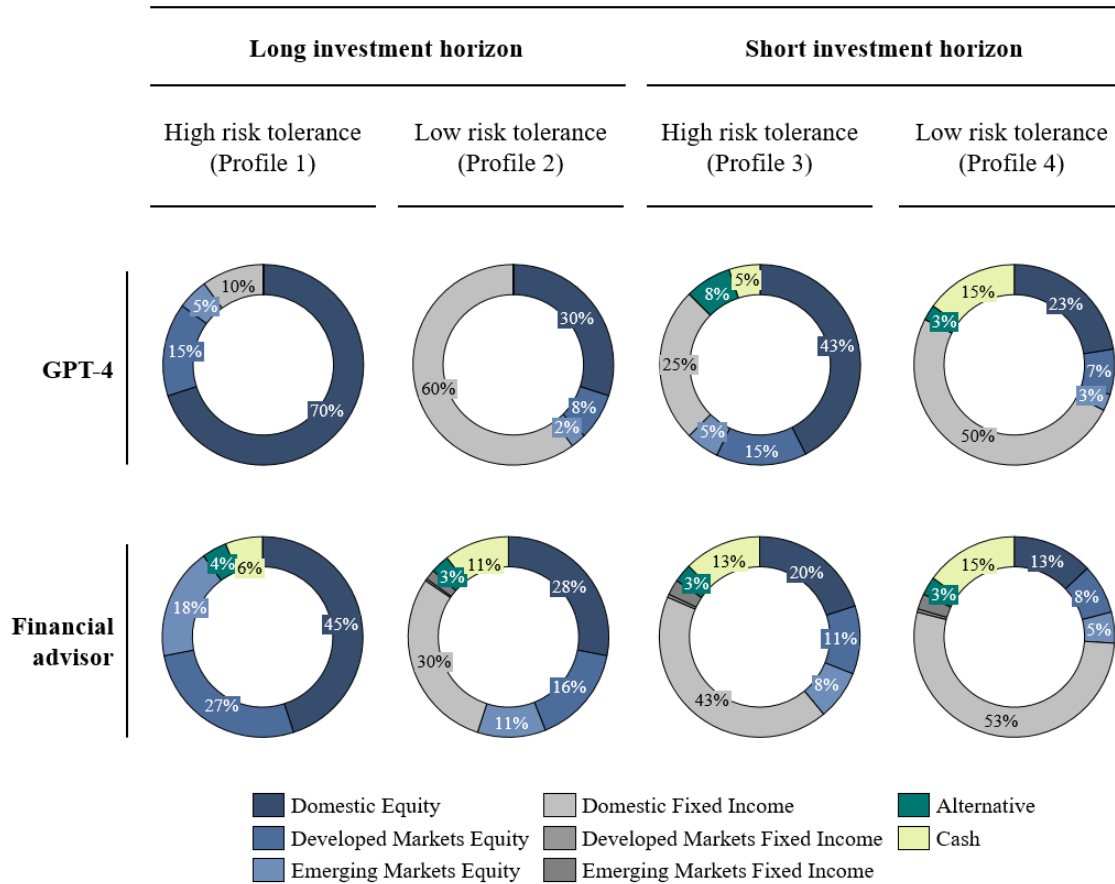
In the next section, we will conduct a more formal evaluation of the quality of GPT's financial advice. As a benchmark for comparison, we elicit portfolio suggestions for our investor profiles from an established professional U.S.-based financial advisory firm's automated financial advice solution, which currently oversees more than \$50 billion in assets under man-

agement. In completing the onboarding questionnaire, we take great care to convey exactly the same information as in the GPT-4 requests (see Table 4 in the Appendix for details).

3 Results

To assess how well diversified GPT’s suggested portfolios are, we compare its composition with respect to geography and asset classes to the benchmark portfolio obtained from the professional financial advisor. We distinguish between domestic (U.S.), developed markets, and emerging markets securities, and equity, fixed income, alternative assets (e.g., real estate or commodities), and cash. Figure 2 displays breakdowns by asset class and geography for each of the four investor profiles.

Figure 2: Portfolio breakdowns



A few things are worth noting here. First, GPT-4 portfolios in general provide exposure to the same geographies and asset classes as professionally advised portfolios. While GPT-4 does

not recommend exposure to foreign fixed income products, these only play a minor role in the benchmark portfolios. Second, GPT-4 portfolios display considerable home bias, both when compared to the benchmark portfolios (for profile 1 80% of total assets and 78% of stocks are domestic vs. 47% of total assets and 50% of stocks) and to the global market capitalization share of U.S. stocks levels (approx. 60%, according to the MSCI ACWI Index⁵, cf. [Bae et al., 2008](#)).⁶ Within international equity, emerging market stocks seem to be particularly underweighted, representing between 5% and 9% of all stocks, compared to approx. 20% in the benchmark portfolios. Third, the portfolio suggestions seem to be more sensitive to risk tolerance and less sensitive to the investment horizon than the benchmark portfolios: While the equity share decreases from 90% to 40%—compared to a decrease from 90% to 55% in the benchmark portfolio—when a long-investment-investor turns risk-averse, it remains at 63% when the investment horizon shrinks to five years—compared to 39% in the benchmark portfolio.

To review the suggested portfolios' performance, we compute monthly average return, volatility figures, and annual Sharpe ratios for the GPT-4 and benchmark portfolios for the period from December 2016 to May 2023, which is the longest time period for which data is available for all investment products in the GPT-4 and benchmark portfolios. Table 2 displays the results, which can be summarized as follows. First, the historical return and volatility figures are consistent with the portfolio breakdowns illustrated in Figure 2. The average returns to GPT-4 portfolios are more sensitive to risk tolerance and less sensitive to the investment horizon than benchmark portfolios. Second, GPT-4 portfolios for the high risk tolerance profiles 1 and 3 have earned significantly higher historical returns than the respective benchmark portfolios ($p < 0.1$, $p < 0.05$). The difference amounts to 26 and 28 basis points per month, respectively. For profile 1, this is achieved without a significant difference in portfolio volatility. For profile 3, portfolio volatility is higher ($p < 0.01$), but the Sharpe ratio is still significantly higher ($p < 0.01$). Third, no differences are found between GPT-4 and benchmark portfolios regarding the historical returns for profiles 2 and 4, while GPT-4

⁵ <https://www.ssga.com/de/de/institutional/etfs/funds/spdr-msci-acwi-ucits-etf-spyy-gy>

⁶ We also observe home bias for a German investor profile in unreported results.

portfolio risk is significantly lower for profile 2 ($p < 0.01$). Taken together, the historical risk and return figures suggest that GPT-4 portfolios provided equal, if not superior risk-return profiles than benchmark portfolios. This finding is also reflected in Figure 3, which shows the evolution of \$100 invested in each of the four GPT-4 and benchmark portfolios over the same period. While both sets of portfolios seem exposed to similar market-wide shocks, such as the Covid-induced dip in early 2020, GPT-4 portfolios have experienced higher growth than their respective benchmark portfolios for all investor profiles.

Finally, to compare portfolio performance after adjusting for common risk factors and to investigate exposures to those risk factors, we estimate the coefficients of a six factor regression model:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i^{Mkt}(R_{m,t} - r_{f,t}) + \beta_i^{SMB} \times SMB_t + \beta_i^{HML} \times HML_t + \beta_i^{RMW} \times RMW_t + \beta_i^{CMA} \times CMA_t + \beta_i^{WML} \times WML_t + \epsilon_{i,t} \quad (1)$$

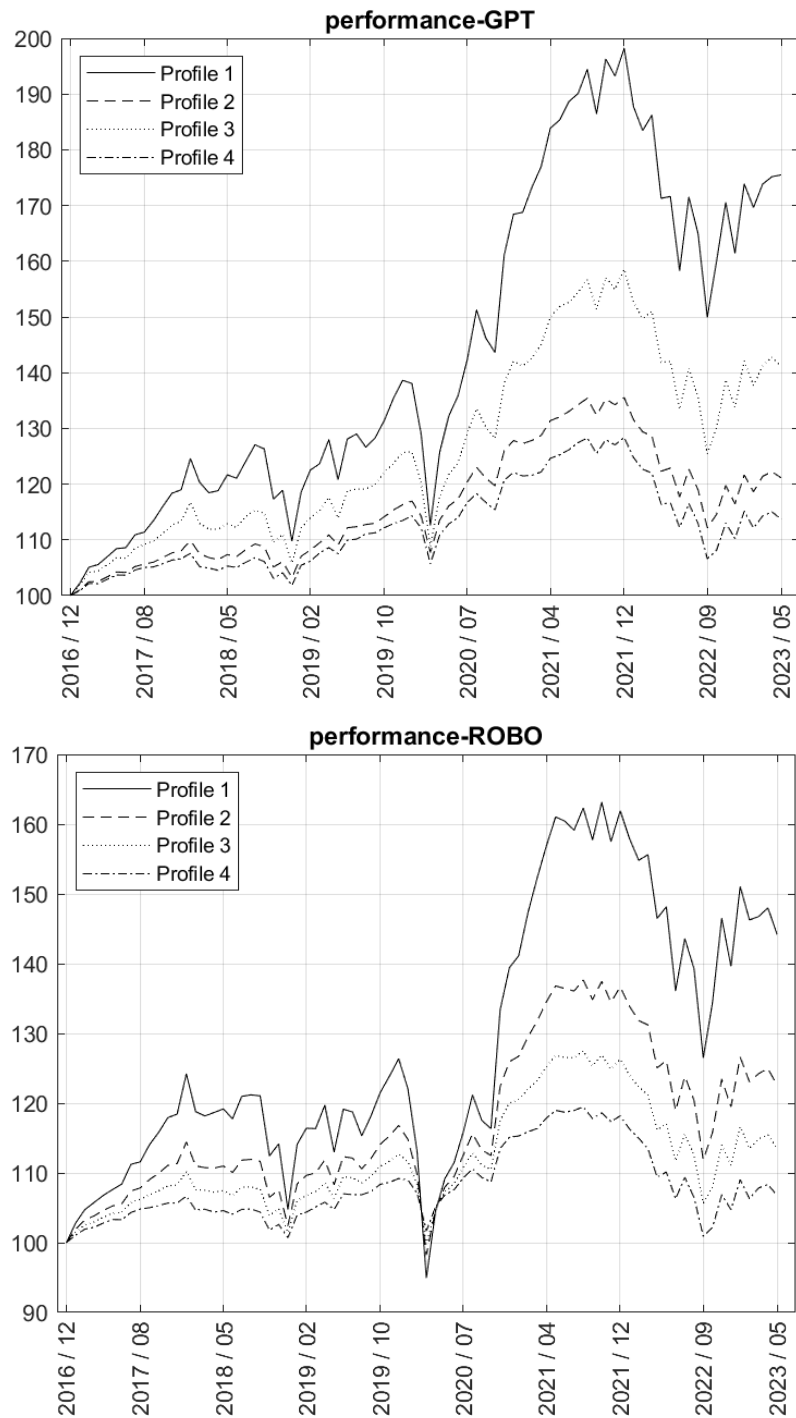
Specifically, we benchmark the excess portfolio returns ($r_{i,t} - r_{f,t}$) against six well-known asset pricing factors: the market excess return ($R_{m,t} - r_{f,t}$), the small minus big (SMB) size factor, the high minus low (HML) value factor, the robust minus weak (RMW) operating profitability factor, the conservative minus aggressive (CMA) investment factor, and the momentum factor—also known as winners minus losers (WML). We use the factors for developed markets and receive them from the website of Kenneth French. Table 3 reports the results of the regression, which can be summarized as follows. First, both GPT-4 and benchmark portfolios have earned negative risk-adjusted returns for profiles 2 through 4. Monthly alphas range from -21 to -26 basis points and are highly significant ($p < 0.01$). For profile 1, the GPT-4 portfolio has earned zero alpha, while the benchmark portfolio has earned a slightly significant negative alpha of 17 basis points ($p < 0.1$). Thus, risk-adjusted performance is nearly identical between the GPT-4 and benchmark portfolios. Second, the market betas confirm the findings in previous sections that GPT-4 portfolios are more responsive to risk tolerance and less responsive to investment horizons than the benchmark portfolios. Third, benchmark portfolios for profiles 1 and 2 (long investment horizons) load positively on the SMB factor, while GPT-4 portfolios for profiles 2 and 3 load slightly negatively, which is

Table 2: Performance metrics

Profile	(1) GPT-4 portfolios			(2) Benchmark portfolios			(1) - (2) Δ				
	N	μ	σ	SR	N	μ	σ	SR	μ	σ	
1	77	0.73	4.54	0.56	77	0.48	4.69	0.35	0.26*	-0.15 (0.78)	0.21 (0.10)
2	77	0.25	2.25	0.38	77	0.26	3.10	0.30	-0.02 (0.91)	-0.85*** (0.01)	0.09 (0.59)
3	77	0.45	3.30	0.47	77	0.17	2.40	0.24	0.28** (0.02)	0.90*** (0.01)	0.23*** (0.01)
4	77	0.17	2.08	0.28	77	0.08	1.79	0.16	0.08 (0.16)	0.29 (0.19)	0.11 (0.21)

Note: This table reports average monthly returns (μ , in %), average monthly volatility (σ , in %), and Sharpe ratios (SR, in %) for portfolios recommended by GPT-4 and the benchmark financial advisory firm for each of the four investor profiles. We test for differences in average returns using a one-sample t test of the return differences, for differences in volatility using a two-sample F test, and for differences in Sharpe ratios using a Ledoit/Wolf test statistic (Ledoit and Wolf, 2008). Significance levels are indicated by stars (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). We draw on data from 2016/12 to 2023/05, which is the longest available time series for which pricing information was available for all recommended products.

Figure 3: Portfolio evolution



Note: This table shows the evolution of \$100 invested in the portfolios suggested by GPT-4 and our benchmark financial advisor for the four investor profiles from 2016/12 to 2023/05.

consistent with a lower share of small-cap funds in GPT-suggested portfolios. Fourth, benchmark portfolios for profiles 1 and 2 load positively on the HML factor, while none of the GPT-4 portfolio does, suggesting a lack of exposure to value stocks in the latter.

Taken together, GPT-4 portfolios have earned comparable risk-adjusted returns to the considered benchmark portfolios, which suggests that the superior performance observed in Table 2 was achieved by exposure to commonly considered risk factors. For profile 3, for which the GPT-4 portfolio displayed higher Sharpe ratio than the benchmark portfolio, this seemed to be mainly due to higher exposure to market risk (market beta of 0.67 vs. 0.46), which is consistent with the considerably larger equity share observed in Figure 2.

4 Discussion and conclusion

The results of our study are substantial. GPT-4 can provide financial advice which is on par with the advice provided by professional low-cost automated financial advisory services. While the portfolios suggested by GPT-4 displayed considerable home bias, its historical risk-return profiles are at least on par with and historical risk-adjusted returns were no worse than benchmark portfolios. Admittedly, the acid test of the suggested portfolios is still outstanding as current recommendations would need to be backtested in 30 years time and the considerable home bias may prove detrimental in the long run. However, any weaknesses in portfolio performance must be traded off against the compounded advisory fee GPT-4 saves investors as well as any improvements in the quality of its responses that are to be expected with future releases.

Aside from the convincing performance of the suggested portfolios, GPT-4 seems to be able to implement recent guidelines for financial advisors: To investigate GPT-4's ability to serve clients' sustainability preferences (ESMA, 2018), we added sustainability preferences to some of our investor profiles. The portfolios suggested for those profiles included ESG-focused versions of the portfolio components such as the iShares ESG Aware MSCI USA ETF.

Finally, while we have shown that GPT-4 is already effective at one part of the financial advisory process—namely matching information on the client's individual situation to a suitable portfolio of financial products—we have not assessed its ability to perform adjacent steps

Table 3: Factor loadings (FF six-factor model)

Profile	N	α (in % per month)	β^{Mkt}	β^{SMB}	β^{HML}	β^{RMW}	β^{CMA}	β^{WML}
<i>Panel A: GPT-4 portfolios</i>								
1	77	-0.10 (0.16)	0.93*** (0.00)	-0.04 (0.37)	0.05 (0.30)	0.01 (0.87)	-0.23*** (0.00)	0.01 (0.67)
2	77	-0.22*** (0.00)	0.43*** (0.00)	-0.10* (0.05)	-0.05 (0.36)	0.04 (0.50)	-0.09 (0.19)	-0.04 (0.19)
3	77	-0.23*** (0.00)	0.67*** (0.00)	-0.10* (0.07)	0.00 (0.93)	0.10 (0.16)	-0.08 (0.31)	0.00 (0.97)
4	77	-0.26*** (0.00)	0.38*** (0.00)	-0.08 (0.21)	-0.04 (0.55)	0.07 (0.35)	-0.10 (0.25)	-0.05 (0.20)
<i>Panel B: Benchmark portfolios</i>								
1	77	-0.17* (0.06)	0.95*** (0.00)	0.18*** (0.00)	0.24*** (0.00)	-0.04 (0.61)	-0.09 (0.30)	-0.04 (0.28)
2	77	-0.21*** (0.01)	0.61*** (0.00)	0.11** (0.04)	0.10* (0.08)	0.00 (0.96)	-0.06 (0.40)	-0.05* (0.09)
3	77	-0.23*** (0.01)	0.46*** (0.00)	0.07 (0.20)	0.04 (0.52)	0.02 (0.82)	-0.05 (0.52)	-0.06* (0.05)
4	77	-0.23*** (0.01)	0.32*** (0.00)	0.03 (0.65)	-0.03 (0.58)	0.02 (0.80)	-0.01 (0.86)	-0.08** (0.02)

Note: This table reports the coefficients obtained from FF-six-factor regressions of excess portfolio returns on the market excess return ($R_{m,t} - r_{f,t}$), the small minus big (SMB) size factor, the high minus low (HML) value factor, the robust minus weak (RMW) operating profitability factor, the conservative minus aggressive (CMA) investment factor, and the momentum factor—also known as winners minus losers (WML). We use the factors for developed markets and receive them from the website of Kenneth French. Standard errors are computed using the method of XXX and are reported in parantheses. Significance levels are indicated by stars (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). We draw on data from 2016/12 to 2023/05, which is the longest available time series for which pricing information was available for all recommended products.

in the advisory process. First, risk profiling, a key part of regulatory guidelines for financial advisors and determinant of financial advice taking (Streich, 2021), can currently not be handled by GPT-4. While risk tolerance is the first factor GPT-4 suggests to consider when making an investment decision⁷ and guiding questions are provided⁸, assigning a risk profile is left to the investor. Second, while specific products, portfolio shares, and exchange tickers are included in the recommendation, GPT-4 cannot offer assistance in implementing the portfolio (opening an account, purchasing and rebalancing portfolio components). Thus, while GPT-4 does well in matching investor profiles to specific portfolios, it will likely not make the entire financial advisory process redundant in the near future. Instead, it may be used as a back-office solution by financial advisors, who use GPT-4 to generate portfolio recommendations from the investor profiles they generate and which they implement and rebalance accordingly.

⁷ Answer to the request “What are important factors to consider when determining a suitable investment portfolio for retirement?”.

⁸ For example, “What are you saving for? Is it retirement, buying a house, starting a business, or a child’s education? Knowing your goals can help determine how much risk you can afford.”

References

- Bae, K.-H., Stulz, R. M., and Tan, H. (2008). Do local analysts know more? a cross-country study of the performance of local analysts and foreign analysts. *Journal of Financial Economics*, 88(3):581–606.
- Calvet, L. E., Campbell, J. Y., and Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, 115(5):707–747.
- Dowling, M. and Lucey, B. (2023). Chatgpt for (finance) research: The bananarama conjecture. *Finance Research Letters*, 53:103662.
- ESMA (2018). Guidelines on certain aspects of the MiFID II suitability requirements 06/11/2018 — ESMA35-43-1163.
- Frey, R., Pedroni, A., Mata, R., Rieskamp, J., and Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances*, 3(10):1–13.
- Jung, D., Dorner, V., Glaser, F., and Morana, S. (2018). Robo-Advisory: Digitalization and Automation of Financial Advisory. *Business and Information Systems Engineering*, 60(1):81–86.
- Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859.
- Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance*, 20(4):587–615.
- Litterscheidt, R. and Streich, D. J. (2020). Financial education and digital asset management: What’s in the black box? *Journal of Behavioral and Experimental Economics*, 87.
- Merton, R. C. . (1969). Lifetime Portfolio Selection under Uncertainty : The Continuous-Time Case. *The Review of Economics and Statistics*, 51(3):247–257.
- Rühr, A., Streich, D., Berger, B., and Hess, T. (2019). A Classification of Decision Automation and Delegation in Digital Investment Management Systems. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*, pages 1435–1444.
- Streich, D. J. (2021). Risk preference elicitation and financial advice taking. *Journal of Behavioral Finance*, pages 1–17.
- The New York Times (2022). The Brilliance and Weirdness of ChatGPT.
- Tobin, J. (1958). Liquidity preference as behavior towards risk. *Review of Economic Studies*, 25(2):65–86.

Appendix

Table 4: Onboarding questionnaire answers, benchmark robo-advisor

#	Question	Answer
1	What is your goal for this account?	Grow investments for future goal
2	What do you want to use this account for?	Saving for retirement
3	What is your understanding of stocks, bonds, and ETFs?	Some
4	When you hear risk related to your finances, what is the first thought that comes to mind?	<i>In line with with risk preferences (highest / lowest)</i>
5	Have you ever experienced a 20% or more decline in the value of your investments in one year?	<i>In line with with risk preferences (highest / lowest)</i>
6	What did you do when you experienced a 20% decline in value of your investments?	<i>In line with with risk preferences (highest / lowest)</i>
7	How would you describe your approach to making important financial decisions?	<i>In line with with risk preferences (highest / lowest)</i>
8	How much do you want to invest to get started?	10k\$
9	How many years from now will you need to start withdrawing funds from this account?	<i>Investment horizon</i>
10	How much do you want to contribute each month?	0\$
11	How much investment value fluctuation would you be comfortable with 1 year from now?	<i>In line with with risk preferences (highest / lowest)</i>
12	Which account type are you interested in?	Taxable
13	Based on the information we've received, you are eligible to enroll in the following portfolio features: Tax-Loss Harvesting, Municipal Bonds	None

Table 5: Full set of profiles

Request	Home country	Age	Investment horizon	Risk tolerance	Sustainability preferences
1	No country	30	40	High	No
2	No country	30	40	High	Yes
3	No country	30	40	Low	No
4	No country	30	40	Low	Yes
5	No country	30	30	High	No
6	No country	30	30	High	Yes
7	No country	30	30	Low	No
8	No country	30	30	Low	Yes
9	No country	45	15	High	No
10	No country	45	15	High	Yes

11	No country	45	15	Low	No
12	No country	45	15	Low	Yes
13	No country	60	5	High	No
14	No country	60	5	High	Yes
15	No country	60	5	Low	No
16	No country	60	5	Low	Yes
17	United States	30	40	High	No
18	United States	30	40	High	Yes
19	United States	30	40	Low	No
20	United States	30	40	Low	Yes
21	United States	30	30	High	No
22	United States	30	30	High	Yes
23	United States	30	30	Low	No
24	United States	30	30	Low	Yes
25	United States	45	15	High	No
26	United States	45	15	High	Yes
27	United States	45	15	Low	No
28	United States	45	15	Low	Yes
29	United States	60	5	High	No
30	United States	60	5	High	Yes
31	United States	60	5	Low	No
32	United States	60	5	Low	Yes
33	Germany	30	40	High	No
34	Germany	30	40	High	Yes
35	Germany	30	40	Low	No
36	Germany	30	40	Low	Yes
37	Germany	30	30	High	No
38	Germany	30	30	High	Yes
39	Germany	30	30	Low	No
40	Germany	30	30	Low	Yes
41	Germany	45	15	High	No
42	Germany	45	15	High	Yes
43	Germany	45	15	Low	No
44	Germany	45	15	Low	Yes
45	Germany	60	5	High	No
46	Germany	60	5	High	Yes
47	Germany	60	5	Low	No
48	Germany	60	5	Low	Yes
