Multi-(Horizon) Factor Investing with AI *

Russ Goyenko[1,2] and Chengyu Zhang[1,2]

[1] *Desautels Faculty of Management, McGill University*
[2] *Financial Innovations and Risk Management Labs, FIRM*

This Version: April 4, 2023

Abstract

Can the backbone technology behind ChatGPT create and manage portfolios? We apply this tech-engine, adapted for finance applications, to multi-factor investing by a long-horizon investor who uses bigger that traditionally used data and takes into consideration long-term versus short-term volatility, liquidity and trading costs trade offs while maximizing expected portfolio returns. The answer is yes, as we are able to actively time factors' premium realizations while dynamically re-balancing and diversifying between factors. Moreover, the long horizon perspective is critical, as it allows for more patient trading and re-balancing needs, more strategic factor timing, and a different set of fundamental signals to rely on.

*We are grateful to Andrea Berardi (discussant), John Cochrane, Stefano Giglio, as well as seminar participants at Durham University, McGill University and participants of 2022 GRETA-CREDIT: Long Run Risks conference (Venice), 2022 Global AI Finance Research Conference (Singapore) for helpfull comments and suggestions.

The authors acknowledge financial support of the Social Sciences and Humanities Research Council of Canada, SSHRC. FIRM Labs: https://firmlabs.ca/.

Corresponding author, Russ Goyenko: ruslan.goyenko@mcgill.ca

Several tech executives and top artificial-intelligence researchers, including Tesla Inc. Chief Executive Officer Elon Musk and AI pioneer Yoshua Bengio, are calling for a pause in the breakneck development of powerful new AI tools.

"We've reached the point where these systems are smart enough that they can be used in ways that are dangerous for society," Mr. Bengio, director of the University of Montreal's Montreal Institute for Learning Algorithms, said in an interview. "And we don't yet understand." ...

The letter doesn't call for all AI development to halt, but urges companies to temporarily stop training systems more powerful than GPT-4, the technology released this month by Microsoft Corp.-backed startup OpenAI. [1]

1 Introduction

A public release of ChatGPT by OpenAI in November 2022 sparked an unprecedented interest and at the same time fear for broad society-wide AI adaptations. In March 2023 main-lead AI developers called for a 6-month pause in AI deployments in order to conduct more research before its further deployment. This paper is a response to this call but in the context of finance applications. More specifically, we provide evidence of whether the back-bone technology behind ChatGPT can be applied to the most basic and fundamental problem in finance - a multi-factor portfolio management.

Why is this important for finance audience? Perhaps it is the best summarized by a different, earlier call in the following quote:

"In sum, I opine that portfolio theory and practice for long-run investors, like all asset pricing, is ripe for important changes. Portfolio theory is perceived as somewhat of a dead end: a 50-year-old model that is so hard to calculate in practice that nobody uses it in the investment world, or a benchmark from which it is easy to accuse even the most sophisticated people and institutions of being behavioral morons." Cochrane (2022)

The choice of multi-factor investing is also not coincidental. While it has been investigated in different contexts by previous literature (see for example Lynch (2001), DeMiguel et al. (2021), Polk et al. (2022)), given the limited number of investable assets (factors)[2] facilitates tractability and exposition of how this new AI technology makes asset allocation decisions.

[1]source: WSJ, "Elon Musk, Other AI Experts Call for Pause in Technology's Development", by Deepa Seetharaman, March 29, 2023

[2]Similar to DeMiguel et al. (2021) we use 9 factors: gross returns for the market (MKT), small-minus-big (SMB), high-minus-low (HML), robust-minus-weak (RMW), and conservative- minus-aggressive (CMA) factors of Fama and French (2015), the momentum (MOM) factor of Carhart (1997), the profitability (ROE) and investment (IA) factors of Hou et al. (2015), and the betting-against-beta (BAB) factor of Frazzini and Pedersen (2014).

In quick summary, it makes optimal portfolio allocation choices exactly the way we model and train it to do, which is fully guided by finance fundamentals. Our fundamental objectives are to address the following questions where traditional techniques currently provide very little guidance: *"How should long-term investors form portfolios? How should they adapt to the fact that our world features time-varying expected returns, volatilities, correlations, and a plethora of factors, signals, and strategies?"* Cochrane (2022).

Hence, we consider long-term investors' optimal strategies who respond to *multiple* factors' realizations, as well as to their future expectations based on factor specific characteristics and macro-economic conditions, as well as the diversification effect between factors in order to form optimal long-term portfolios. Our modelling objective is, therefore, two-fold. First, we condition portfolio weights allocation among asset classes on "plethora of factors, signals, and strategies" or the big, or bigger that traditionally used data.[3]

Second, all short- and long-term investors are cautious about portfolio turnover and trading costs, as ignoring them results in large utility losses (Balduzzi and Lynch (1999)). Moreover, long-term investors, who have longer time horizon to reach their target portfolio are expected to be even more cautious about the current turnover if they can spread the rebalancing needs over multiple months ahead. Therefore, we want our model, before it makes asset allocation decisions for out-of-sample investment period, to recognize and be explicitly trained on the multi-horizon investment perspective. We set ten years, 120 months, as the maximum investment horizon, and 1 year, 12 months, and the minimum, with explicit penalties on assets turnover and trading costs for each rebalancing time period while maximizing portfolio expected returns and minimizing its risk for the whole investment horizon. In the main tests we allow for monthly rebalancing across all investment horizons. We also consider quarterly/semiannual/annual re-balancing for 3/5/10 year investment horizons respectively.

We find that short term, 12 month horizon strategy is more impatient, has higher portfolio turnover and volatility compared to long, 120 month investment horizon strategy. The monthly portfolio turnover of short term investor is 11%, while it is only 6% per month for long term portfolio, i.e. almost 50% difference. Moreover, portfolio turnover monotonically decreases with the investment horizon, while the Sharpe ratios monotonically increase from 2.344 for short term to 2.74 for long term horizon. The nine factor risk adjusted monthly portfolio returns, alphas, are virtually the same across all investment horizons, 51 bps per month (6.12% annualized) with t-statistics exceeding 7. The difference in the Sharpe ratios is attributed to substantially lower volatility of long-term portfolio.

[3] We use 153 firm specific characteristics from Jensen et al. (2022) to construct 153 factor-specific features. Besides these features for each factor we also use a wide range of macro-economic variables aimed to capture business cycle, market and funding liquidity, and overall macro-economic hidden states.

Among lower re-balancing frequencies we obtain the Sharpe ratio of 3.051 for semi-annual re-balancing and 5 year investment horizon. Interestingly, for annual re-balancing and 10 year investment horizon we obtain almost similar Sharpe ratio, 2.702, compared to monthly re-balancing for the same investment horizon.

The long-horizon strategy is more patient in trading and timing factor realizations as it has ten times more, 12 vs 120, future time intervals to strategically spread its trades. The shorter horizon, on the other hand, causes more aggressive trading every month, and as a result more portfolio volatility. We also find an empirical support to Polk et al. (2022) as our long horizon investor allocates twice more to the value, HML, factor, and holds much lower position in momentum, UMD, compared to the short-term investor, which is consistent with predictable mean reversion in returns argument. This mean reversion is only observed for longer horizons, and empirically for HML, and not for UMD which is of a short-term nature. Consistent with the same argument, we also find that the long-term portfolio invests 3 times more in SMB on average, compared to the short term portfolio.

Furthermore, we find that re-balancing frequency and investment horizon are important determinants of the primary information set of variables the model relies on. Analyzing the variable/feature importance, we find that short term, 12 months, investment horizon consideration uses co-skewnes, skewness, betas, return volatility, idiosyncratic skewness or reversals as the top predictors, which is also similar to the results in the previous literature (Gu et al. (2020, 2019)). For longer investment horizons, 36 to 120 months, and monthly re-balancing, we observe the mix with variables reflecting future cash flow expectations, such as labour force efficiency, earning persistence, net operating assets, or free cash flow-to-price ratio appearing in the top 5 most important predictors.

The biggest contrast is observed for the lower re-balancing frequencies. Here, we find that earnings surprises, 1 year sales growth, changes in sales to inventory ratios, free cash flow to price, earning persistence, net-operating assets, tax expense surprise, Ebitda-to-market enterprise value are the top predictors for lower (quarterly/semi-annual/annual) re-balancing frequencies which take into account long, up to 10 years, investment horizon period considerations. Moreover, none of top 20 important variables determining the success of long term strategies overlaps with those of 12 month investment horizon and monthly re-balancing. An important message here is that even a relatively passive strategy with annual re-balancing, but long-horizon perspective, by relying on a different information set, can perform as well as short-term, relatively active strategy with monthly re-balancing frequencies. This in turn has applications for ETFs construction for example which are not-actively managed portfolios.

How do we contribute to the literature? First, we support the literature which opposes the idea that long-term investing should be passive buy-and-hold strategy. For example, Campbell and Viceira (1999) argue that, when expected asset returns and hence investment opportunities are not constant, an optimal portfolio policy of long term investor should involve timing of the stock market and hedging in order to avoid large welfare losses. More recently, Moreira and Muir (2019) analyze a portfolio choice problem of long horizon investor in a partial equilibrium model when both expected returns and volatility are allowed to vary over time. In contrast to conventional view that long-term investors are better off ignoring temporary movements in volatility, the authors show significant costs from ignoring volatility timing. DeMiguel et al. (2021) show the benefits of Moreira and Muir (2017) volatility timing in a multi-factor portfolio context for short, one month investment horizon. We go further and show multi-factor timing and its lager benefits for long horizon investors.

Second, our results clearly point towards factor timing ability which we achieve with the newer technologies and using more data. This finding supports Haddad et al. (2020) who find that 50+ factors constructed to reflect various anomaly characteristics are predictable and hence factor timing strategies can be profitable. We go further and show that fundamental factors' timing, while controlling for diversification effect among factors on a portfolio level, can lead to further improvements in timing strategies.

Third, and related to a bigger picture, we argue that any portfolio of a risk averse investor should have a long term perspective. Majority of the current empirical literature and industry practices consider one month ahead perspective. We show that this short-term perspective, which is 12 months in our setting, makes the identification of optimal portfolio weights rely on short-term signals related to time series return characteristics. It leads to higher turnover and overall volatility on a portfolio level. A long term perspective disciplines portfolio weights allocations to rely on long-term fundamental signals, which leads to higher capital preservation, i.e. lower draw-down, during higher volatility regimes - the ultimate objective of any risk-averse investor.

Fourth, we can provide an answer to the most asked question after the release of ChatGPT - can it manage or create portfolios/ETFs? We see some big financial companies to start moving into this space. For example, the news article about Citadel initiating enterprise-wide licence for ChatGPT with OpenAI mentions: *"ChatGPT potentially could create stock portfolios, or hasten the production of analyst presentations. There's even an exchange-traded fund planned around the concept."*[4] Our results clearly show that not ChatGPT per se, but the technology behind ChatGPT definitely can, once it is adapted towards finance fundamental

[4]Bloomberg News, March 07, 2023: "Citadel Negotiating Enterprise-Wide ChatGPT-License" https://www.bnnbloomberg.ca/citadel-negotiating-enterprise-wide-chatgpt-license-griffin-says-1.1892360).

principles.

What kind of fundamentals we rely on while building the AI model architecture? First we admit the fact that predicting returns is a difficult task, and especially in a multi-horizon setting. Instead of predicting future returns, we predict portfolio weights allocations. That is instead of training the model to predict returns, we train a trading strategy, i.e., dynamic optimal portfolio with implicit return forecasting. Signals useful for the trading strategy are not always the most useful signals for pure return forecasting. This approach allows incorporating various "limits to arbitrage" considerations by design.

Second, we also use the fact that long-horizon returns are better predicted by fundamentals, like dividend-price ratio, DP, compared to short-horizon returns (Fama and French (1988)), and especially annual returns (Golez and Koudijs (2018)). To demonstrate a basic portfolio policy rule we consider a dynamic asset allocation strategy just between two factors, MKT and BAB (betting agains beta, Frazzini and Pedersen (2014)), a small set of predictive asset specific characteristics (size, book-to-market and momentum, similar to Brandt et al. (2009)), and only one state variable, dividend-price ratio, DP (Campbell and Viceira (1999)). By construction BAB is market neutral and hence can be considered as a hedge portfolio, and DP has been shown to predict market returns better at longer horizons (Fama and French (1988), Campbell and Viceira (1999), Brandt (1999), Golez and Koudijs (2018)). When DP is excluded, there is no signal for market timing strategies, as we know that size, book-to-market or momentum do not predict MKT. As a result, the model almost does not trade (the turnover is only 2.5% for monthly or 5.7% for annual rebalancing), and passively allocates almost equal weights between MKT and BAB. That is the model is not given any information or market timing signal that we know works. When DP is included as the state variable, the model starts actively trading, the monthly turnover increases to 14.3% (33.3%) for monthly (annual) re-balancing and 12 month (10 year) investment horizon. The trading is consistent with market timing, as the model ex-ante allocates higher weights to factors when DP ex-post realizations are in the lowest percentiles, i.e. the model times the factors correctly (Brandt (1999), Campbell and Viceira (1999)). Given that DP predicts MKT the best at the annual frequencies (Golez and Koudijs (2018)), we obtain the highest Sharpe ratio for annual re-balancing and 10 year horizon, 1.16, compared to the Sharpe ratio of 0.58 for the monthly re-balancing and 12-month investment horizon. Furthermore, the only significant both economically and statistically risk-adjusted portfolio alpha is observed for the annual re-balancing. Better predictability of annual returns (Golez and Koudijs (2018)) by DP allows for better market timing.

Third, our final exercise follows Gu et al. (2020) to allow for more data and signals to time the factors, and we allow for more than two factors, similar to Lynch (2001), from the current factor pricing literature (Fama and French (2015), and Hou et al. (2015)) to be able to achieve better diversification while spanning the most of fundamental risks.

As a special case, we analyze how the model allocates the weights around COVID-19 March, 2020, market plunge. 2020 is the last year of our *OOS* period, and the model was retrained last time at the end of 2019. Moreover, the model has never been trained on anything similar to this pandemic episode as it never happens before in our sample period, 01/1980 to 12/2020. This event therefore provides a unique laboratory experiment to examine how after being trained on previous crisis episodes, the model makes decision for something it has never experienced before. Consistent with the predictions of Moreira and Muir (2019), we find that the model decreases the weight of MKT before March 2020 when the market volatility started increasing, and then increases it in the end of March, i.e. it advised to buy MKT at the market bottom. The model also advised shorting SMB, i.e. taking negative exposure to small cap, at the end of January 2020 and covering most of short position in March 2020, when this short-position was the most in-the-money. These ex-ante model decisions can be found quite rational and effective ex-post by long-term institutional, i.e. pension funds, portfolio managers who cannot short-sell the market, but can time the market volatility (Moreira and Muir (2019)) or reduce their positions in small cap stocks when volatility is high.

The rest of the paper is organized as follows. Section 2 describes related literature. Section 3 discusses our technology choices. Section 4 outlines the base technology adaptation for financial applications, model architecture and overall methodology. Section 5 describes the data, and out-of-sample tests' results. Section 6 discusses economic fundamentals determining the success of long horizon portfolio investment strategies. Section 7 concludes the paper.

2 Related Literature

Our modelling approach follows the literature which draws inferences about optimal portfolio weights without explicitly modeling the underlying return distributions. Brandt (1999), Aït-sahali and Brandt (2001) and Brandt and Santa-Clara (2006) model the optimal weights between three asset classes as non-parametric functions of several state variables, i.e. dividend yield, term and default premiums, that predict assets' returns. Brandt et al. (2009) model the portfolio weights of individual stocks as a linear function of three stock characteristics: size, book-to-market and momentum. More recently Chen et al. (2020) and Chatigny et al. (2021) model portfolio weights of individual stocks as non-linear functions of multiple stock and macro-economic features.

Our approach is also similar to Lynch (2001), as we consider portfolio allocation by multi-horizon investor among multiple factors. Unlike Lynch (2001), who primarily focuses on size and book-to-market characteristic portfolios, we consider all currently available factors capturing different risk exposures (Fama and French (2015), Hou et al. (2015)). Further, Lynch (2001) assumes a dynamic model relating returns to two, dividend yield and term spread, forecasting variables and solves for an investor's portfolio and consumption choice using estimates of the implied conditional distribution of returns. An incorrect model of how returns relate to forecasting variables can yield invalid inferences (Brandt (1999)). Moreover, the recent literature argues that there are many more other variables that can predict stock returns (Gu et al. (2020)). Importantly, Lynch (2001) does not incorporate transaction costs into the analysis, while Balduzzi and Lynch (1999) demonstrate significant utility losses from ignoring them in a multi-period setting.

We also relate to more recent literature arguing in favor of factor timing and factor-returns predictability (Haddad et al. (2020)), and the other extensive literature showing that incorporating stock return predictability into asset allocation choices improves overall portfolio performance (Brennan et al. (1997), Balduzzi and Lynch (1999), Brandt (1999) Barberis (2000), Lynch (2001)).

Another important literature explores how portfolio performance, rebalancing strategies and asset allocations depend on the investment horizon. Prior studies analyzing long horizon strategies consider asset allocation between stock index and cash (Brennan et al. (1997), Barberis (2000), Campbell and Viceira (1999)). They find that when dividend-price ratio positively predicts market's returns, the market's Sharpe ratio increases with horizon and long-horizon investors allocate higher weights to stocks than short-horizon investors. In contrast we consider long horizon investor allocating between multiple risk factors. The closest to our paper is Polk et al. (2022) who study dynamic portfolio choice of two factor portfolio, value and momentum, by long term investor in a calibrated equilibrium model. The authors find that because of predictable long horizon mean reversion in the value factor, and relatively short memory of momentum, a long horizon investor allocates more weight to the value, and less to momentum compared to short horizon investor. None of the above papers considers dynamic turnover and rebalancing strategies of short vs long horizon investors while allocating across multiple risk factors.

3 Technology Choices: Discussion

Our modelling objective is two-fold. First, we condition portfolio weights allocation among asset classes/factors based on a large number of predictors. We use 153 firm specific characteristics from Jensen et al. (2022) to construct 153 factor-specific features. Besides these features for each factor we also use a wide range of macro-economic variables and allow these variables to inter-act with factor-specific features, similar to Gu et al. (2020).

Second, we want our model to explicitly recognize long-term investment horizon. This implies different trade offs between rebalancing and turnover strategies for short vs. long horizons, as well as different risk management strategies, while maximizing expected portfolio pay offs.

Standard modelling techniques used in the classical asset pricing literature cannot accommodate the long-term dynamically forward looking conditional expectations of time varying factor returns, volatility, illiquidity and macro-conditions, as well as simultaneously accounting for diversification effect across factors on the portfolio level in one consistent framework. They also can not account for dynamic turnover strategies across multiple investment horizons. Yet, these objectives are the key to success of any long-term multi-factor asset allocation strategy.

This is a high-dimensional problem which otherwise could not be addresses without more recent technological developments. Machine learning, ML, is the best approach to handle high-dimensional problems. However, a commonly used supervised learning to long term investment horizon strategy cannot really be applied. Supervised learning is learning from a training sample of labeled examples provided by an expert. In the current ML asset pricing literature expected stock returns are often used as labeled targets for training (Gu et al. (2020)). For a long term investor, in our setting, a training objective is to optimize a 10 year optimal portfolio performance path, before we make an asset allocation decision for the first out-of-sample month. This path should optimally time factor realizations over 120 consecutive months, or 10 years, accounting for co-variations of this timing with other factors in the portfolio, rather than striking the best possible performance in the first year. This optimal path cannot really be labeled, but it can be learned by trying out various scenarios to identify which one is the best to maximize a certain numerical reward, which can simply be the mean-variance utility net of transaction costs of a risk averse investor.

Reinforcement learning, RL, a type of unsupervised learning, involves learning what to do in order to maximize a given numerical objective. The learner here is not told which action to take as for example in a supervised learning to minimize return forecast errors (Gu et al. (2020)). Instead, an algorithm must discover which actions yield the most reward by trying them out. The most interesting part which makes RL most applicable for long term investment strategy optimization is that actions affect not only the immediate reward but also the next situation/scenario, and through that all subsequent rewards. Therefore, consequences of actions, and how their short- vs- long-term trade offs play out for the final long term realizations are the key features of RL. This allows to train the model in a such way that, for example, the optimal asset allocation in year 1 takes into account all possible consecutive rebalancing needs for years 2, 3, 4, ... and 10, in order to "spread" rebalancing as well as factors' timing over years rather then accomplishing it all at once in one period. Economically, the long term objective allows for example accommodating a conventional 5-year business cycle window and learning about efficient asset allocation choices conditioned on the latent state of macro-cycles. RL is to some extent a deep learning, non-parametric, modern-day alternative to dynamic programming used by the previous literature (Brennan et al. (1997), Lynch (2001))[5] in high dimensional space.[6] In finance literature RL was first successfully used by Cong et al. (2021) for long-short alpha portfolio strategy.

3.1 Market Efficiency

We choose to train our model similar to ALphaGo, via RL, and let the algorithm to determine the most optimal path towards maximizing its objective. AlphaGo has never been taught by humans, and it ended up beating the best of them. This approach is based on learning relentlessly to beat humans in their own game.

Translating it towards asset management applications - do we know what the optimal portfolio is, or what it can be? We only obtain the optimal portfolio based on the techniques we have, and those are often based on the assumptions, linearity, short-term investment horizon, and using a limited information set. Most of the traditional techniques do not allow using big data.

As a result we often infer optimal portfolio either based on the modelling restrictions, or observing the best portfolio managers. The question is - do these best portfolio managers make the best possible investment decisions. Or do the models we have allow for the best obtainable empirical output? Both are affected by human choices and assumptions. Technically, we never let the data to speak freely for themselves. The principle of AlphaGo, or an idea to beat the best portfolio managers in their own investment game, can speak directly to the market efficiency. If we do not fully use the data, or do not fully allow the data to speak for themselves - the markets will never be fully efficient.

3.2 Technology

Figure 1 presents the evolution of technology leading up to ChatGPT. GPT is Generative Pre-trained Transformer, the model architecture developed by OpenAI. The transformer

[5]One important difference from classical dynamic programming approach is that instead of modeling the end outcome in the future and *back-ward* solving to the present for intermediate horizons it solves *forward* for the future *expected* multi-period outcomes to find optimal solutions for intermediate steps.

[6]RL has been successfully applied in robot control, AlphaGo, or self-driving cars.

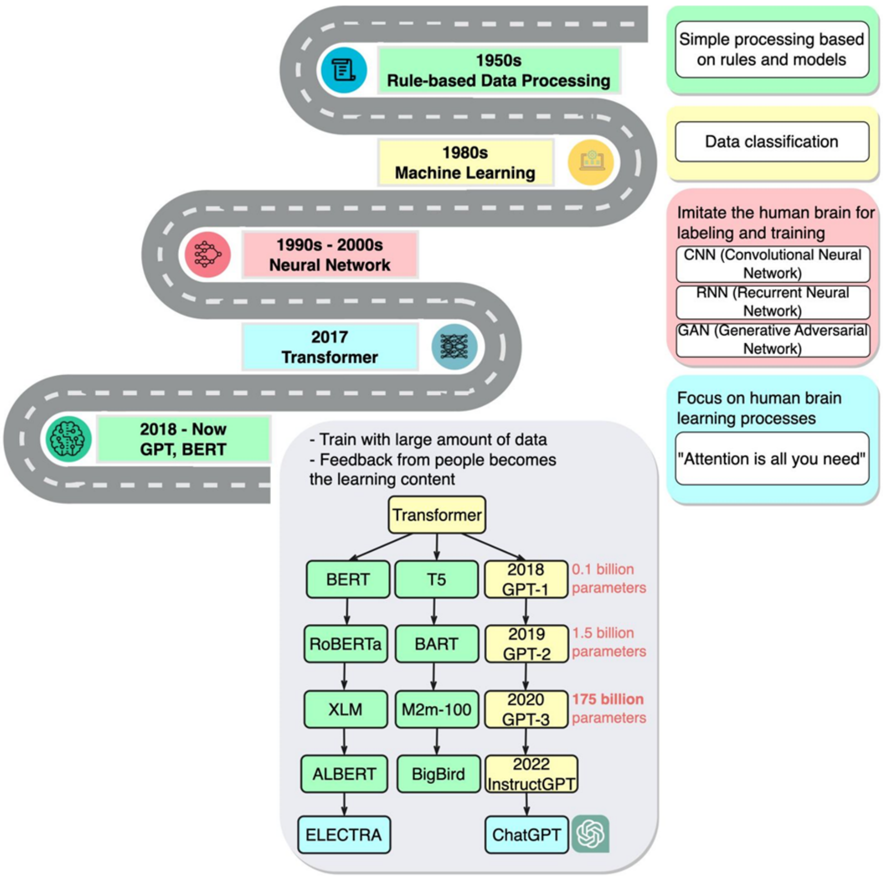


Figure 1: Technology Evolution and ChatGPT

architecture was first introduced by Vaswani et al. (2017) and it relies on self-attention mechanisms. It immediately replaced the recurrent, convolutional and generative neural networks due to its much higher efficiency and significantly improved inferences for language models. It also serves as a foundation for Large Language Models, LLMs, such as BERT, and GPT. It is a back-bone engine behind ChatGPT, and perhaps the most powerful learning network yet that we have ever seen. It evolved NLPs towards LLMs, where the "bag of words" approach has been replaced by contextual understanding of the text. That is what allows ChatGPT to be so "intelligent".

The question is - can this technology create an intelligent portfolio manager? Deep and shallow learning with feed-forward neural networks, FNNs, or GANs (Generative Adversarial Networks) have been extensively discussed in the empirical asset pricing literature (Gu et al. (2020); Bianchi et al. (2021); Chen et al. (2020)). While FNNs are now widely used in finance literature, they have only cross-sectional or panel data applications, and cannot account for time series structure of the data. Moreover their predictive ability and out-of-sample performance have been mostly attributed to identifying anomalies as its practical applications can easily run into the limits-to-arbitrage problems (Avramov et al. (2021)). Finally, these Neural Networks more than a decade predate the more recent and advanced ones based on Transformer, which uses a completely different, attention-based signal extraction mechanism.

The off-shelf Transformer is a very complex network, which comes with encoder and decoder used for machine language translations. The complexity of this network is simply not applicable for finance data as it would require, by some computations, at least 1000 years of monthly data for training, which we do not have. Instead we adapt, and much simplify the core of this network to be able to apply for financial markets' data. In the next section we describe the key architecture choices to keep it to the minimum level of complexity possible yet retaining the most powerful features of this network.

4 Model and Methodology

One of the key algorithms in our model architecture is Transformer Encoder (TE). TE is a part of sequence representation extraction models (SREM), which till recently was dominated by Long Short-Term Memory (LSTM). Both, TE and LSTM allow to extract information from the time series of high dimensional input features. The difference is that LSTM belongs to the class of recurrent neural networks (RNNs), and as the name suggests it processes the input data sequentially based on their time-stamps. To process information in month t for example, the model combines all available information from previous months leading up to month t. Therefore, information from say month t-12 can propagate arbitrarily down the sequence all the way to month t, which from financial economics perspective, for example, can bias us to identifying mostly momentum driven signals. Unlike RNNs, and LSTM in particular, TE processes the entire input all at once, without modeling sequential dependencies in the time series. Given the low auto-correlation structure of financial data, e.g. stock returns, TE is perhaps the most suitable for finance applications among all other currently available ML approaches. It takes all time series input at once, and instead of trying to model its time series sequential dependence, which does not exist in the most of financial data, it allocates different attention weights according to a learned measure of relevance and assigns a relevant importance to any input regardless where it is time-stamped positioned in the time series. This is particularly valuable for long-term asset allocation. For example, the last month firm's cash flows would not be the most valuable information for its stock future 10 year price growth. Moreover, trying to model sequential dependence between four consecutive earnings announcements over the last 12 month, i.e. LSTM approach, can be viewed as fruitless exercises. However, different pieces of information collected about this firm during

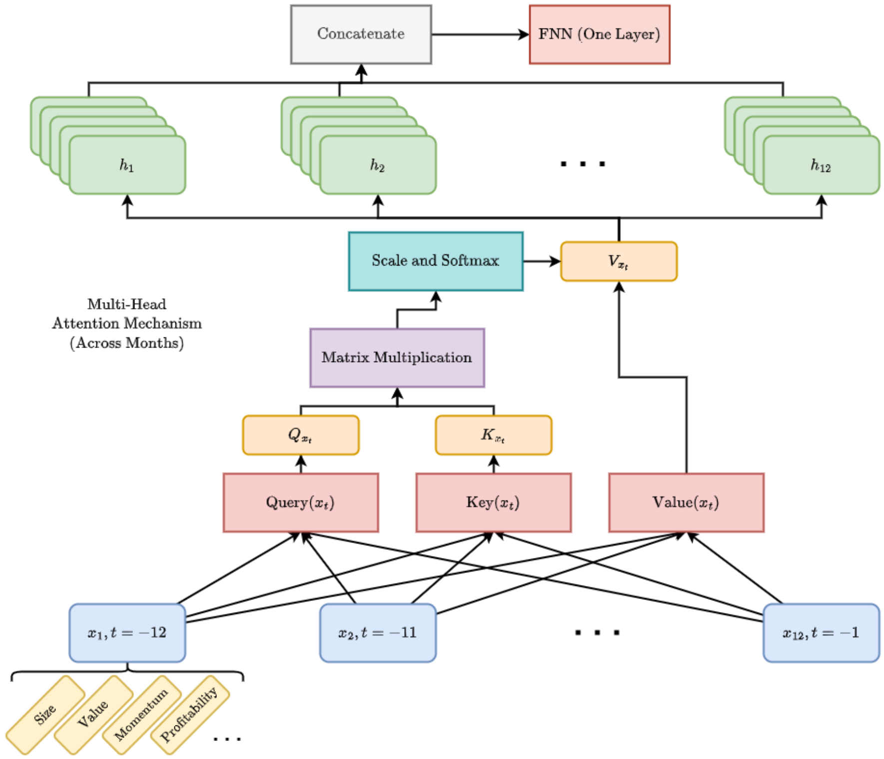


Figure 2: Transformer Encoder Architecture

last 12 months, e.g. different four quarters' earnings surprises and weighing them differently for long term growth identification, can provide richer and more complete economic picture.

The measure of relevant importance, which is the key component of TE and allows differentially weighting the significance of any part of the input data, is based on self-attention mechanism (Vaswani et al. (2017)). This mechanism is often used in NLP to identify the similarities among words in the sentence. This similarities are aimed to best describe the object in the sentence. We use attention mechanism to extract the most relevant representations and signals from time series and cross-section of multiple asset characteristics which best describe and contribute to the best future portfolio performance. Our TE is different from regular TE in NLP applications.

Figure 2 presents our basic TE architecture for one asset. The input here, x_t, are 12 vectors associated with 12 month of look-back window. Each vector contains 153 asset specific characteristic for each month.

First, TE performs sparsity reduction of each vector by reducing its dimensionality to 128. After that, a reduced vector of each month characteristics undergoes embedding with queries, Q_{x_t} and keys, K_{x_t}. Queries is a set of vectors that we want to calculate attention for. Keys is a set of vectors that we want to calculate attention against. As a result of dot product multiplication, we obtain set of weights a (also vectors) showing how attended each Query against Keys.

For example, assume that one of the elements in Q_{x_t} is size, and we want to understand how attended, or relevant the size characteristics versus for example other elements in K_{x_t}, such as momentum, book-to-market, volatility and etc. for the future portfolio performance. The matrix multiplication between Q_{x_t} and K_{x_t} provide a vector with the set of weights, where size has an importance weight relative to all other characteristics. Similarly, all other characteristics have the relative weights against non-overlapping counter-parts. This vector then goes through softmax standardization to assure that the weights in the end add up to 1. These weights are probabilities of relevance or importance of the input characteristics for objective function maximization. After that, this vector of weights is multiplied by the vector of values,V_{x_t}, for each month. As a result, all embedded asset specific characteristics, after their dimensionality has been shrunk from 153 to 128 for each month, are weighted by their importance. Subsequently, these 12 vectors are concatenated at the level of the first attention head, h_1.

Overall we use multi-head attention with four heads. This means that the Attention module repeats its computations four times in parallel. It is quite an advantage when it comes to the textual analysis. This means that separate four sections of the Embedding can learn different aspects of the meanings of each word, as it relates to other words in the sequence. This allows the Transformer to capture richer interpretations of the text overall.

For our purposes, one can argue that one head is enough. We are not reading the text where the words can have different contextual properties and hence all possible interpretations should be considered. There are only so many ways that size, for example, can dominate book-to-market or momentum in order of relevant importance. We however continue using three other heads for robustness, as there are no really extra costs - these are parallel processes.

The output of all four heads is further concatenated into one vector which then goes as an input into FNN with one layer, the last block at the top of Figure 2. This linear layer performs the final sparsity reduction providing the final output of 12x153 which is similar to the original input dimension. This output is different as it contains month-characteristic weighted features which are the most important across all four heads for the portfolio optimization. That is all monthly information of a look-back window, i.e. a matrix of 12x153 for each asset, gets first truncated and shrunk based on the importance of each month-characteristic to the future portfolio performance, and then the remaining part is weighted with the probability weights based on each variable-month relevant contribution to the future portfolio performance across four attention heads.

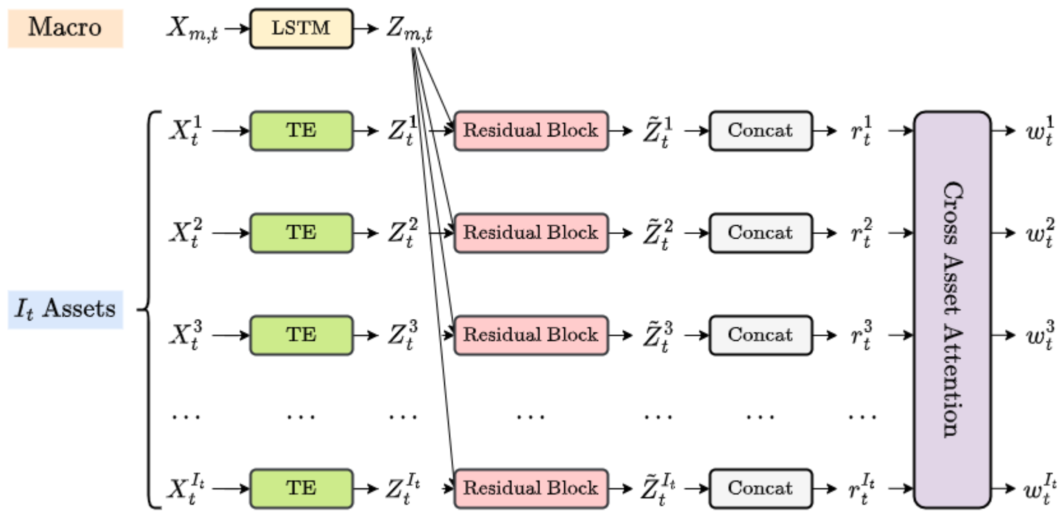


Figure 3: Full Model Architecture

While TE is the key block of our model, it is not the full structure of the model yet. The final model contains nine TEs: one TE for each of nine assets in our portfolio. Further, the importance of adding and conditioning portfolio optimization on the latent state of macro-economy has been extensively discussed by Chen et al. (2020). Here, similar to Chen et al. (2020), LSTM is a natural choice to extract the latent macro-state from 12 month look-back window, as the time series and path-dependence of macro-processes, which LSTM is perfectly suited to capture, define business cycles. Where we are currently in the business cycle and its future expectations are an important input while forming a long-term portfolio. Moreover, Zhang and Aaraba (2022) demonstrate in the context of bottom-up portfolio construction and similar model architecture that adding macro-economic latent state information to portfolio optimization is crucial for superior portfolio performance in high volatility market regimes.

We subsequently add FNN with one layer to allow the output of LSTM, the latent state of macro-economy, to inter-act with the output from TE for each asset, the weighted variable-months which contribute the most to better portfolio performance.

Figure 3 illustrates the overall architecture of the model. To bring the information from different blocks of the model together, let us denote a vector $\tilde{\boldsymbol{x}}_t^{(i)}$ as a state history of asset i at time t which consists of asset specific features/characteristics, and similarly a vectors of $\tilde{\boldsymbol{x}}_{m,t}$ as a vector of macro-economic variables at time t. We define the last K historical holding periods at time t, i.e., the period from time $t-K$ to time t, as a look-back window of t. For both, $\tilde{\boldsymbol{x}}_t^{(i)}$ and $\tilde{\boldsymbol{x}}_{m,t}$, in all estimations we use 12-month look-back window to predict, for example, asset allocation weights for the month 13th.

The historical states from the look-back window of asset-specific information as well as

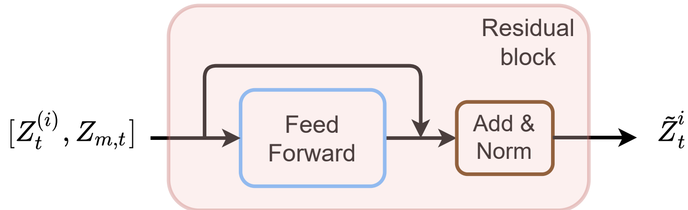


Figure 4: Residual Block

macro indicators are then denoted respectively by the sequences $X_t^i = \{x_1'^i, \ldots, x_j'^i, \ldots, x_K'^i\}$ and $X_{m,t} = \{x'_{m,1}, \ldots, x'_{m,K}\}$, where $x_j'^i = x_{t-K+j}^i$ and $x'_{m,j} = x_{m,t-K+j}$.

The Transformer Encoder, TE, (Vaswani et al. (2017)) projects the asset-specific sequence X_t^i into a latent vector space as

$$Z_t^i = \mathrm{TE}(X_t^i), \tag{1}$$

where $Z^i = \{z_1^i, \ldots, z_K^i\}$. Each z_t^i represents the hidden state encoded at step k, and in the case of Transformer Encoders, also contains information from all other hidden states. Subsequently, LSTM (Hochreiter and Schmidhuber (1997)) extracts macro indicators latent representation,

$$Z_{m,t} = \mathrm{LSTM}(X_{m,t}), \tag{2}$$

where $Z_{m,t} = \{z_{m,1}, \ldots, z_{m,K}\}$.

After the latent states have been extracted, we feed the concatenation of the representation of each asset i and the macro latent state, $[Z_t^i, Z_{m,t}]$, to a residual block which consists of a fully connected feed-forward network, FNN, followed with residual connection and layer normalization as shown in Figure 4. We use a simple default specification for this FNN with 1 layer and 512 neurons. The output of this block, $\tilde{Z}_t^i = \{\tilde{z}_t^1, \ldots, \tilde{z}_t^K\}$, is then concatenated to form a one dimensional vector representation of the corresponding asset: $r_t^i = \mathrm{Concat}(\tilde{z}_t^1, \ldots, \tilde{z}_t^K)$, which is a description of the asset based on asset-specific characteristics, macro latent states at time t, as well as interaction effects between the two.

Finally, the last block of the model consists of the Cross Asset Attention Network (CAAN). CAAN is based on self-attention mechanism of the Transformer's architecture discussed above in Figure 2. Cong et al. (2021) modify this attention mechanism from being asset specific centered, self-attention, to allow capturing across-assets dependencies, cross-asset attention, contributing to the final optimization objective. That is given a representation r_t^i learned for an asset i at time t, CAAN allows estimating portfolio weights w_t^i for the asset i depending on its relative importance compared to all other assets in the portfolio.

Following slef-attention mechanism, for each asset i, we compute vectors $\boldsymbol{q}^i, \boldsymbol{k}^i, \boldsymbol{v}^i \in \mathbb{R}^d$, representing query, key and value vectors respectively, with subscript t being removed to ease the equations:

$$\boldsymbol{q}^i = \boldsymbol{W}_Q \boldsymbol{r}^i, \quad \boldsymbol{k}^i = \boldsymbol{W}_K \boldsymbol{r}^i, \quad \boldsymbol{v}^i = \boldsymbol{W}_V \boldsymbol{r}^i, \tag{3}$$

where $\boldsymbol{W}_Q$, $\boldsymbol{W}_K$, and $\boldsymbol{W}_V$ are asset-independent parameter matrices to be optimized. Subsequently, the rescaled cross-asset relationship (Cong et al. (2021)) between asset j and asset i is described with $\boldsymbol{q}^i$ to query the key $\boldsymbol{k}^j$ as

$$\beta_{ij} = \frac{\boldsymbol{q}^{i\intercal} \cdot \boldsymbol{k}^j}{\sqrt{d_k}}. \tag{4}$$

where d_k is a re-scale parameter which is essentially the variance of the dot product to prevent it from becoming too large.

In other words, CAAN, instead of using Q_{x_t} and K_{x_t} of one individual asset as discussed in Figure 2, uses Q_{x_t} of one asset against K_{x_t} of all other investable assets.

We futher define the normalized cross-asset relationships, β_{ij}, as the weight:

$$\beta_{ij}^* = \frac{\exp(\beta_{ij})}{\sum_{j=1}^{I} \exp(\beta_{ij})}, \tag{5}$$

We then use these weights to sum up the values $\{\boldsymbol{v}^i\}$ of all other assets into an attention vector:

$$\boldsymbol{a}^i = \sum_{j=1}^{I} \beta_{ij}^* \cdot \boldsymbol{v}^i. \tag{6}$$

From this point we compute portfolio weights through a fully-connected layer applied to the attention vector $\boldsymbol{a}^i$ as $w^i = W_c \cdot \boldsymbol{a}^i + e_c$, with W_c and e_c being the weight and bias parameters to learn.

4.1 Training via Reinforcement Learning and Objective Function

We train our model via RL. In our RL training, an agent interacts with the environment in a T-period investment game. The environment consists of a series of state-action-reward trajectories, i.e. $\pi = \{\text{state}_1, \text{action}_1, \text{reward}_1, \ldots, \text{state}_T, \text{action}_T, \text{reward}_T\}$, where $state_t$ is the historical market state characterized by historical information described by the two tensors $\mathcal{X}_t = \{X_t^i, i = 1, \ldots, I_t\}$ and $\mathcal{X}_{m,t}$. $action_t$ is agent's action given the state, $state_t$, where $action_t^{(i)}$ represents the optimal weight of an asset in the portfolio. $reward_t$ is the reward given action_t, which could be the optimal portfolio return for example. Thus, a trajectory denotes the interactions between the agent and the market environment throughout a fixed period of time defined by the investment horizon T. We call such trajectory an episode.

Let H_{π_θ} be an objective function of the portfolio manager. Then its input is the trajectory of rewards, $\{\text{reward}_1, \ldots, \text{reward}_T\}$, and in the end of each trajectory (episode), the trajectory reward of the agent is described as the average reward acquired during the investment horizon $H_{\pi_\theta} = \frac{1}{T}\sum_{t=1}^{T} \text{reward}_t$.

The reinforcement learning optimization algorithm finds the optimal parameters $\theta^* = \text{argmax}_\theta H_{\pi_\theta}$ through stochastic gradient ascent, i.e. $\theta_t = \theta_{t-1} + \eta \nabla H_{\pi_\theta}|_{\theta=\theta_{t-1}}$, where η is the learning rate, and ∇H_{π_θ} is automatically computed through back-propagation. We employ the Adam Optimization algorithm of Kingma and Ba (2015) to automatically shrink the learning rate towards zero as the gradient approaches zero.

When the model is empirically trained, an episode is defined as one, three, five or ten years of investment horizon T for the monthly, quarterly, semi-annual and annual rebalancing respectively. The monthly rebalancing uses T of 12 months and monthly returns, quarterly rebalancing uses T of 12 quarters, i.e. 3 years, and quarterly returns, semi-annual rebalancing uses T of 10 semi-annual periods, i.e. 5 years, and semi-annual returns, and finally annual rebalancing uses T of 10, 10 years, and annual returns.

When we implement monthly portfolio rebalancing across all investment horizons, the definition of T is much simpler: 12, 36, 60 and 120 months for 1, 3, 5 and 10 year horizons respectively, and we use monthly returns through out.

The output of the optimization process are portfolio allocation weights which can be defined in the following general form for each individual asset i, and time t:

$$\omega_{t,i} = f\left(Z_t^i, Z_{m,t}\right) \tag{7}$$

Relaying it to RL framework, and H_{π_θ} objective function, the final functional form of weights is:

$$\tilde{\omega}_{t,i} = f_\theta\left(Z_t^i, Z_{m,t}\right) \tag{8}$$

Our optimization problem is optimal multifactor portfolio based on optimal holdings from $t-1$ to t, $\tilde{\omega_{t-1}}$, that optimize the mean variance utility net of transaction costs of an investor with risk aversion parameter γ over holding period T:

$$\max_{\omega_{t-1}} \mathrm{E}\left[\sum_{t=1}^{T} \rho^{t-1}\left(\omega_{t-1}^\top \mu_t - \frac{1}{2}\gamma\omega_{t-1}^\top \Sigma_t \omega_{t-1} - \frac{1}{2}\lambda \Delta\omega_{t-1}{}^\top c_t \Delta\omega_{t-1}\right)\right], \tag{9}$$

where μ_t and c_t are the realized return and trading costs in period t respectively.[7] Σ_t is the covariance matrix of returns for the period t estimated differently for each rebalancing

[7]In other wards the investor has mean-variance preferences over the change in her wealth W_t each time period, net of risk free return: $\Delta W_{t+1} - r^f W_t = \omega_t^\top \mu_{t+1} - c_{t+1}$.

frequency. For monthly frequency it is computed using 12-month look-back window of returns plus month t, the 13th return. For the quarterly rebalancing, we use 12-month look-back window of monthly returns plus monthly returns of the quarter t, i.e. 15 monthly returns overall, and so on for other lower rebalancing frequencies. Further, while the realized trading costs, c_t, for monthly rebalancing are just the average trading costs for the month t, for the quarterly rebalancing, c_t are realized monthly average trading costs in the last month of the quarter t, which corresponds to the time when the actual rebalancing occurs, and similarly for other lower frequency rebalancing periods. For monthly rebalancing across all investment horizons we use the approach similar to the 12 month horizon explained above. Note that we use realized returns and trading costs as forward looking information only for the training purposes, i.e. only in the training sample, which is common in RL training. Our out-of-sample investment forecasts, as we describe below, after the model being trained, only use look-back window historical data to make predictive asset allocation choices.

$\rho \in (0, 1]$ is a discount factor, γ is the agent's risk-aversion, and λ is the penalty coefficient for illiquidity.[8]

It is important to note that every $action_t$ has a long-lasting propagating effect for each trajectory, as for example $action_{t-1}$ affects $action_t$ via $\Delta\omega_t = \omega_t - \omega_{t-1}$, turnover, and associated with it trading costs, c_t.

Our objective function assumes quadratic transaction costs which is motivated by the linear price impact models. Shutting down the trading cost and turnover penalty by setting λ to 0, the optimal portfolio to hold for each period is similar to the conditional Markowitz portfolio $M_t = (\gamma\Sigma_t)^{-1}\mu_t$. However, when trading costs or other market frictions are present, it becomes optimal for the agent to deviate from the conditional Markowitz portfolio by taking possible future realizations of trajectories into account.

5 Empirical Analysis

5.1 Data and Training Sample Empirical Design

Our asset classes are risk factors commonly used in the literature: gross returns of the market (MKT), small-minus-big (SMB), high-minus-low (HML), robust-minus-weak (RMW), and conservative- minus-aggressive (CMA) factors of Fama and French (2015), the momentum (MOM) factor of Carhart (1997), the profitability (ROE) and investment (IA) factors of Hou et al. (2015), and the betting-against-beta (BAB) factor of Frazzini and Pedersen (2014). These are also the exact set of factors used by DeMiguel et al. (2021) in their multifactor

[8] Without loss of generality, we assume $\rho = 1, \gamma = 1, \lambda = 1$ in all our estimations.

Panel A. Summary Statistics

Factor	Mean	Std. Dev.	Skewness	Kurtosis	Sharpe
MKT	0.0071	0.0439	-0.5939	1.2894	0.563
SMB	0.0015	0.0306	0.3773	4.5641	0.167
HML	0.0005	0.0307	0.0875	2.8023	0.053
RMW	0.0032	0.0259	-0.4699	11.3286	0.429
CMA	0.0017	0.0201	0.6022	1.9744	0.288
UMD	0.0049	0.0477	-1.4359	10.3890	0.354
IA	0.0016	0.0200	0.3583	1.6575	0.285
ROE	0.0042	0.0277	-1.0165	5.9613	0.521
BAB	0.0077	0.0384	-0.3994	3.2995	0.693

Panel B. Correlation Matrix

	MKT	SMB	HML	RMW	CMA	UMD	IA	ROE
SMB	0.247							
HML	-0.106	-0.041						
RMW	-0.358	-0.449	0.359					
CMA	-0.357	-0.046	0.628	0.222				
UMD	-0.283	-0.051	-0.227	0.071	0.012			
IA	-0.317	-0.129	0.643	0.293	0.909	-0.013		
ROE	-0.435	-0.444	0.071	0.686	0.107	0.504	0.153	
BAB	-0.276	-0.109	0.398	0.478	0.349	0.260	0.369	0.417

Notes: The table presents monthly factors' returns summary statistics and correlation matrix for the whole sample, 1980 to 2020. The factors are: gross returns of the market (MKT), small-minus-big (SMB), high-minus-low (HML), robust-minus-weak (RMW), and conservative- minus-aggressive (CMA) factors of Fama and French (2015), the momentum (MOM) factor of Carhart (1997), the profitability (ROE) and investment (IA) factors of Hou et al. (2015), and the betting-against-beta (BAB) factor of Frazzini and Pedersen (2014).

Table 1: Factors' Returns Summary Statistics and Correlations, Full Sample

conditional mean-variance portfolio construction with the market volatility timing.

We download the factors from the corresponding authors' websites, and then replicate these nine value-weighted portfolios in order to estimate: (i) transaction costs required to trade the stocks comprising these portfolios; (ii) portfolio-levels characteristics given the stocks in these portfolios.

Table 1 presents summary statistics of the monthly factor returns, Panel A, and their correlation matrix, Panel B, for the whole sample period. Consistent with stylized facts, the factors' distributions are far from normal, with majority of them having negative skewness, and RMW and UMD having the highest tails measured by kurtosis. MKT and BAB have the highest annualized Sharpe ratios of 0.563 and 0.693 respectively, although they are negatively correlated, -0.276, Panel B. Further, RMW and ROE have one of the highest positive correlation among factors, 0.643, while IA and CMA have the highest, 0.91, correlation.

Our monthly data cover the historical period from January 1980 to December 2020. To estimate transaction level trading costs, the high-frequency trading data, TAQ, are not available for the whole sample. We therefore use the measure of bid-ask spread proposed by Corwin and Schultz (2012), and since the measure provides an estimate of round-trip, i.e. buy and sell, percentage bid-ask spreads, we divide it by 2 to capture half-spread associated with factor portfolio rebalancing, i.e. either buy or sell.

The data on individual stock characteristics are from Jensen et al. (2022) and comprise of 153 individual monthly, or quarterly in case of some COMPUSTAT data, stock-specific characteristics.[9] Table 1 in Appendix summarizes these variables. We retain NYSE/AMEX/NASDAQ common stocks which have at least 85% of these stock characteristics available across the whole historical sample. Using these stocks we are able to replicate the above 9 factor portfolios with 0.90 and higher correlation with the original factors obtained from the corresponding authors websites.

For each factor portfolio we construct 153 factor characteristics by applying the same weighting scheme to individual stock characteristics that are used for stock returns to construct the factors.[10] The only exception is trading costs where long or short positions always incur positive trading costs, and thus the weights are always positive.

Besides 153 features that we construct for each factor portfolio, we also use the following macro-economic variables: 12 macroeconomic series following the variable definitions detailed in Welch and Goyal (2008), including dividend-price ratio (dp), dividend yield (dy), earnings-price ratio (ep), stock variance (svar), book-to-market ratio (bm), net equity expansion (ntis), Treasury-bill rate (tbl), long term rate of returns (ltr), term spread (tms), default spread (dfy), default return spread (dfr), and Consumer Price Index (infl)[11]. To the set of these variables we also add the returns on S&P500 index, as well as market wide illiquidity measured by Amihud (2002) ILLIQ ratio (amihud) and bid-ask spread measure of Corwin and Schultz (2012) after bottom-up aggregation of the estimates at the stock level. To proxy for the market-wide funding illiquidity we use Frazzini and Pedersen (2014) BAB factor. Overall we have 16 market-wide and macro-economic indicators.

Training. To train the model, we use the data from January 1980 to December 2004 as our first training sample. Before the training begins, model parameters are randomly initialized. To start the training for the monthly rebalancing, we randomly draw a month from the training set without replacement. We build historical states using stock characteristics and

[9]We are grateful to Bryan Kelly for making these data available.

[10]Following standard practice in the literature, stock-level characteristics are cross-sectionally ranked and mapped into the [-1,1] interval before they are aggregated up to form factor-level characteristics.

[11]The monthly data for these variables are from Amit Goyal's website

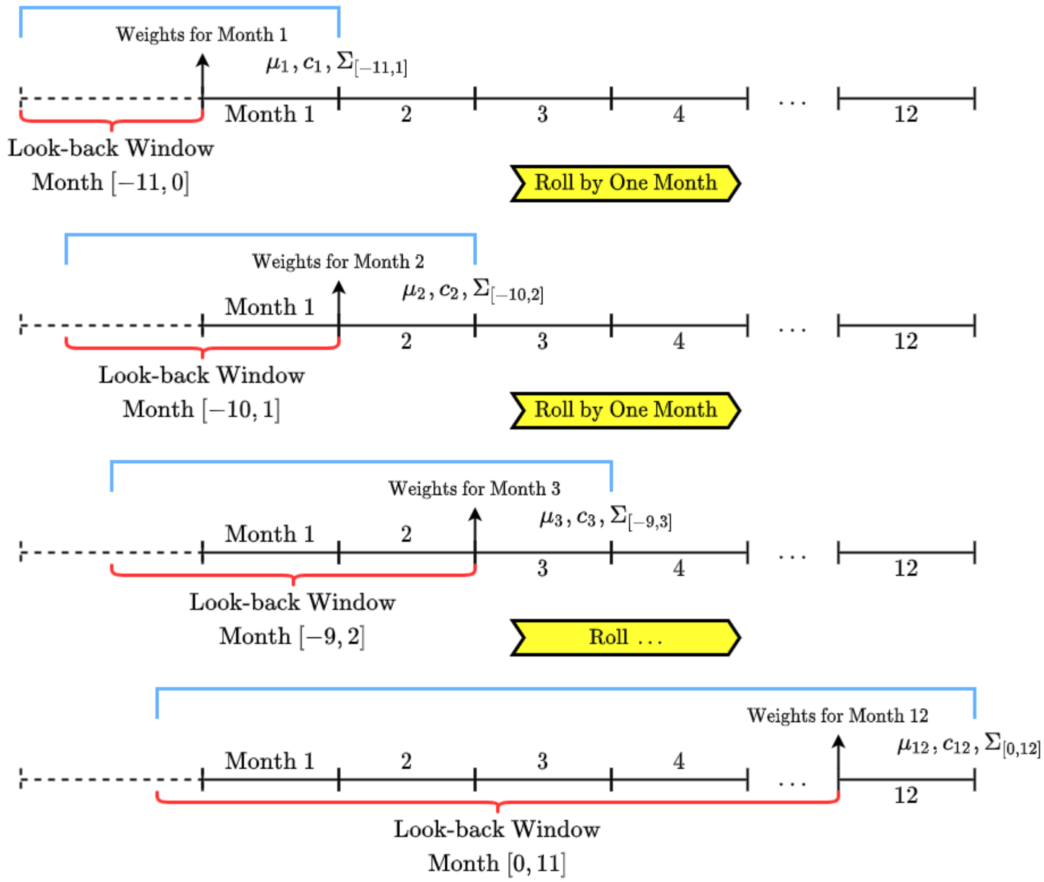


Figure 5: Training RL Model: Monthly Rebalancing

macroeconomic variables from the preceding 12 months. These historical states are processed by the model which outputs the portfolio weights for the selected, 13th, month. We obtain the weights for the subsequent 11 months using the same procedure, each with corresponding 12-month historical states. Essentially, we randomly draw a unique 24-month window from the training set.

Figure 5 visualizes this procedure for the case of monthly rebalancing and 12 month investment horizon. Here, the Months 1 to 12 define an episode, T, or the holding period horizon, and Month 1 is the first randomly selected month in the training sample we mention above. The top panel depicts the initiation of the optimization. The procedure first estimates the vector of weights for Month 1 using all historical information from 12-month look-back window, $[-11, 0]$, the realized returns in Month 1, μ_1, the average trading costs for Month 1, c_1, and the returns from the look back window plus the returns from Month 1 to estimate covariance matrix, $\Sigma_{[-11,1]}$. Then it rolls the estimation window by 1 month, Figure 5 the second panel, to estimate the vector of portfolio weights for Month 2, or the 14th month for the current 24 month window. The look back window now includes Month 1 information, $[-10, 1]$, and we are using the realized returns from Month 2, μ_2, the average trading costs for the Month 2, c_2, and the most recent 13 monthly returns to estimate $\Sigma_{[-10,2]}$ for Month 2. We continue this rolling procedure all the way till Month 12 to finalize the current episode.

Once we obtain 12 monthly sets of portfolio weights, we optimize the objective function in eq. 9 and the model parameters accordingly. We repeatedly draw from the training set until we exhaust all the available overlapping unique 24-month windows, and optimize through all episodes. This multi-episode process that uses all the data in the current training sample is called an epoch. Overall, we train the model with a maximum of 100 epochs. To avoid overfitting the model, we adopt "early stopping" by keeping track of the objective values. If the value of the objective function does not increase for five consecutive epochs, the training process is terminated, and the model parameters from the epoch with the highest objective value are saved.

To avoid high computational cost of updating the deep reinforcement models at high-frequency, i.e. monthly, we refit the model annually.[12] Our first *OOS* testing period is thus from January 2005 to December 2005 with monthly rebalancing. After the model is trained, we use the last look-back window of 12 months, January 2004 to December 2004, to make the optimal weights predictions for January 2005. We invest in the beginning of January 2005, and hold this portfolio through the end of January, and record realized returns. We then roll the 12-month look-back window by 1 month, which now includes January 2005, keep the model parameters from the first training sample, and thus the only update is the macro- and asset specific features which just have been recorded for the month of January. We than estimate optimal weights allocation for the month of February 2005, invest in the beginning of February, hold the portfolio through the end of February, and record realized returns, and so on. Once we reach December 2005 to make the weights allocation decision for January 2006, we retrain/refit the model by rolling the training sample forward by one year while keeping its 25-year length fixed, and so on. Overall, our *OOS* period is from January 2005 to December 2020.

For monthly rebalancing and longer investment horizons, 36 to 120 months, we apply the exact same approach as for 12 month investment horizon. Instead of 24-month window, for 36 month horizon, for example, we draw a 48-month window. We use the same rolling procedure as above and instead of 12 monthly sets of portfolio weights we obtain 36 to finalize one episode, and then we optimize the objective function. The longest computational time is required for 120 month investment horizon, where we first obtain 120 sets of weights to

[12]This is a standard practice in the literature, see for example Gu et al. (2020) and Cong et al. (2021).

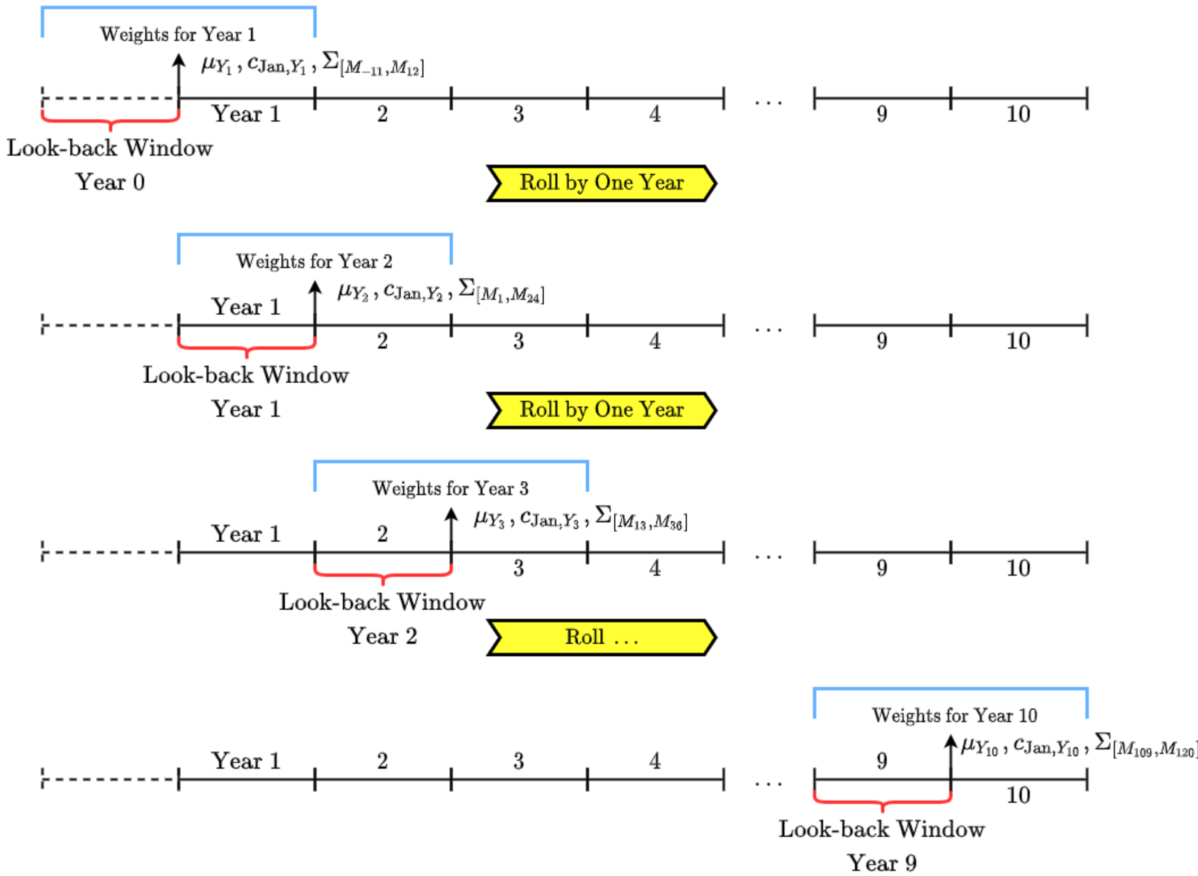


Figure 6: Training RL Model: Annual Rebalancing

optimize the objective function.

Our next focus is on lower, less aggressive, frequency of rebalancing, and still long-term investment perspectives. Using similar approach as for the monthly rebalancing, Figure 5, we examine quarterly, semi-annual or annual portfolio rebalancing. The episodes' lengths T in training are still 36 months, 60 months and 120 months for quarterly, semi-annual and annual rebalancing, respectively, while the look-back window always remains 12 months. Our OOS testing period is still one year, and the model is refitted every year regardless of the rebalancing frequency.

For example, consider annual rebalancing, which, given the descriptions above, entails 10 year holding period horizon considerations, i.e. the length T of an episode is 120 months or 10 years for which monthly returns are cumulated to the annual. Adding to it the 12 months look-back window results in a random draw of 132 month window from the training sample.

Figure 6 demonstrates the training procedure for the annual rebalancing. Similar to

Figure 5, the look-back window to estimate annual portfolio weights for Year 1 remains 12 months, or the historical monthly data for Year 0. For the training purposes, we also use the realized, annual, return in Year 1, μ_{Y_1}, the average realized trading costs for January of Year 1, c_{Jan,Y_1}, and the most recent 24 months of returns as for the end of Year 1 to estimate $\Sigma_{[M_{-11},M_{12}]}$, which we annualize afterwards to be on the same scale as μ_{Y_1}. The average trading costs c_{Jan,Y_1} imply that the portfolio rebalancing, and the realization of trading expenses will occur once a year, in the beginning of the calendar year.

Once we obtain the first set of weights for Year 1, the estimation is then rolled by one year forward, the second panel, Figure 6. The new look back window is now 12 monthly historical periods of Year 1, and we obtain the weights for the Year 2 using the end of Year 2 realized annual returns, μ_{Y_2} , January of Year 2 average trading costs, c_{Jan,Y_2}, and the end of Year 2 most recent 24 monthly returns to estimate $\Sigma_{[M_1,M_{24}]}$ which we annualize. Here, again, the average trading costs c_{Jan,Y_2} imply that the portfolio rebalancing, and the realization of trading expenses occurs once a year, in the beginning of Year 2. We continue this rolling procedure all the way till the end of Year 10 to finalize the current episode.

Once we obtain 10 annual sets of portfolio weights, we optimize the objective function in eq. 9 and the model parameters accordingly. We repeatedly draw from the training set until we exhaust all the available overlapping unique 132-month windows, and optimize through all episodes. As before, this multi-episode process that uses all the data in the training set is called an epoch, and we train the model with a maximum of 100 epochs, and all other details remains the same as before.

Therefore, the model is trained to provide efficient estimates for 10 consecutive annual portfolio holding vectors of weights. Then, similar to the monthly rebalancing approach, we make the first *OOS* prediction for the year 2005 using as a look-back window the monthly data from January 2004 to December 2004, and so on.

5.2 Main Results: Monthly Re-balancing for Different Investment Horizons

Table 2 reports *OOS*, January 2005 to December 2020, RL portfolio performance statistics for monthly rebalancing and different investment horizons. The Alphas are estimated using two models: (i) Fama and French (2015) five factor plus Carhart (1997) momentum, 6-factor model; (ii) 9-factor model which incorporates all 9 factor portfolios we use in our analysis and has the following representation:

	12m	**36m**	**60m**	**120m**
Return	0.0072	0.0069	0.0067	0.0065
Std.Dev.	0.0106	0.0089	0.0086	0.0082
Sharpe	2.344	2.686	2.689	2.741
6 Factor Alpha	0.0053	0.0056	0.0057	0.0057
6 Factor Alpha T	8.254	8.900	9.080	9.272
6 Factor IR	2.133	2.462	2.455	2.544
9 Factor Alpha	0.0051	0.0050	0.0051	0.0051
9 Factor Alpha T	7.636	8.377	8.616	9.021
9 Factor IR	2.086	2.407	2.442	2.569
MaxDD	0.044	0.034	0.027	0.033
Max 1 Period Loss	-0.032	-0.021	-0.017	-0.016

Notes: The table reports *OOS* RL portfolio performance statistics and turnovers for an investors with monthly rebalancing and different investment horizons, ranging from 12 months (12m) to 36 months (36m), 60 months (60m) or 120 months (120m). All statistics are monthly except Sharpe ratios and IR which are annualized. 6-factor Alpha and IR (information/appraisal ratio) are estimated from Fama-French 5 factor model plus momentum factor, while 9-factor model includes all 9 factor portfolios in our analysis. IR is computed as Alpha divided by residual standard deviation after regressing portfolio returns on either 6 or 9 factors, and is annualized. The *OOS* period is from January 2005 to December 2020.

Table 2: *OOS* RL Portfolio Performance - Monthly Rebalancing, and Multi-horizon investment Perspectives

$$P_t = \alpha + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 RMW_t + \beta_5 CMA_t + \beta_6 UMD_t + \beta_7 IA_t + \beta_8 ROE_t + \beta_9 BAB_t + \varepsilon_t \quad (10)$$

where P_t is the time series of our multifactor portfolio returns. IR $= \alpha/\sigma(\varepsilon_t)$ and can be interpreted as the Sharpe ratio when all the factors on the right hand side are hedged out (i.e., the alpha expressed via Sharpe ratio). Sharpe ration and IR are annualized. All other numbers are in monthly format.

One important observation is that the Sharpe ratios are monotonically increasing with the horizon, from 2.344 for 12 month to 2.741 for 120 month investment horizon. The monthly portfolio returns and especially 9 factor Alphas are very similar across all investment horizons. The difference comes in portfolio volatility monotonic decrease with the investment horizon. Here, a short term investor, with 12 month horizon, is more aggressive in timing the market

	12m	36m	60m	120m
Avg TC × Turnover All	0.0009	0.0007	0.0005	0.0005
Turnover All	0.110	0.078	0.065	0.061
Turnover Long	0.078	0.058	0.047	0.042
Turnover Short	0.032	0.020	0.017	0.018

Notes: The table reports *OOS* monthly turnovers for an investors with monthly rebalancing and different investment horizons, ranging from 12 months (12m) to 36 months (36m), 60 months (60m) or 120 months (120m). The *OOS* period is from January 2005 to December 2020.

Table 3: *OOS* RL Portfolio Performance - Monthly Rebalancing, and Multi-horizon investment Perspectives - Turnover

and has the highest monthly portfolio standard deviation of 1.06%. The long-term investor, with 120 month investment horizon, is more patient, and has the monthly standard deviation of 82 bps, almost 23% lower.

The last two rows of Table 2 report portfolio risk statistics: MaxDD, the maximum draw-down, i.e. the maximum consecutive loss from one period to another, and the maximum 1 period loss, Max 1 Period Loss. As the investment horizon increases, these, conservative to begin with, statistics become smaller. For comparison, the MaxDD for S&P500 in our *OOS* period is 55%, while it is only 3.3% for 120 month investment horizon.

Consider, for example, 12 month horizon. The portfolio average return is 72 bps per month, and 6- and 9-factor Alphas are 53 bps and 51 bps respectively, both highly statistically significant after adjusting standard errors with Newey-West correction for auto-correlation and heteroscedasticity with 3 lags. 51 bps alpha obtained with 9-factor model results in 6.12% annualized risk-adjusted return. It shows that our portfolio, which uses exactly the same factor investment as 9-factor model, significantly outperforms all static combinations of these 9 factors by economically significant magnitudes.

Even more important evidence of horizon effect on portfolio timing strategies is observed for turnover, Table 3. Here, TC is the average portfolio level measure of one way transaction relative bid-ask effective spread, where bid-ask spreads are estimated using Corwin and Schultz (2012) methodology. Turnover is estimated as the average change in weights across all holding periods. Even though all portfolios are rebalanced monthly, the monthly turnover monotonically decreases with the horizon. A short-term portfolio with 12 month investment horizon has monthly turnover of 11%. The monthly turnover of long, 120 months, horizon investors is almost half of it, 6.1%. Given that a long-term portfolio has many more rebalancing periods compared to short-term portfolio, it allows the former to time the market more strategically, and spread extreme rebalancing needs over multiple months. This not only reduces portfolio turnover almost by 50%, but also portfolio volatility, and contributes to overall outperformance by long-term investor.

Long horizon portfolio even further outperforms a short-term, 12 month horizon portfolio after trading costs. The average monthly turnover costs for 12 month portfolio are 9 bps, or 108 bps per year. For Long-term portfolio these costs are 5 bps per month or 60 bps per year. Therefore, after trading costs, long-term portfolio gains are significantly higher.

We next consider how asset allocation weights depend on the investment horizon. There is little guidance in the theoretical literature on how long vs short term investors should allocate across multiple risky factors. Recently, Polk et al. (2022) consider dynamic portfolio choice between value and momentum factors and theoretically argue that optimal momentum portfolio weights relative to value decline significantly as horizon increases. That is longer horizon investors prefer to invest more in value, and less in momentum portfolio compared to short horizon investors. The intuition is that momentum has short memory based on most recent performance and its returns are approximately independent over time when evaluated over longer horizons. In contrast, value has long memory because it goes long in assets which underperformed over long time. This increases vaue's expected returns and results in strong negative long-horizon autocorrelation of value returns.

Polk et al. (2022) consider portfolio of these two factors and dynamic weights allocations between them based on both diversification effect and their return predictability. The authors also find significant contribution of diversification to the ultimate improvement in the portfolio performance. We have a portfolio of 9 risk factors, with significantly larger room for diversification effect. We therefore cannot quantitatively directly relate our results to those in Polk et al. (2022). Qualitatively, however, we have very similar mechanism in place, and hence should observe more value investing by long horizon investor.

Table 4 compares the average factor weights' statistics between two extremes, a short-term investor with 12 month investment horizon versus long term investor with 120 month horizon. The performance of these strategies is reported in Table 2, with monthly portfolio rebalancing for both. Long term investor does invest substantially more in value, HML, then short term investor. The average weight of HML for long horizon investor, 8%, twice exceeds the average weight invested in HML by the short term investor, 4.1%.

Applying the same intuition to SMB as to HML, that small stocks can be undervalued for a long time which makes their expected returns higher in the long run, we should expect long term investors to invest on average more to SMB. Indeed, our long term investor invests almost three times more on average, 5.9%, into SMB, compared to the short-term investor, 2.1%.

UMD has been underperforming over our *OOS* period. Both investors short sell UMD.

		12m					120m		
Factor	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Factor**	**Mean**	**Std Dev**	**Minimum**	**Maximum**
BAB	0.124	0.096	-0.159	0.296	**BAB**	0.119	0.067	-0.017	0.323
CMA	0.014	0.060	-0.114	0.173	**CMA**	0.000	0.041	-0.106	0.066
HML	**0.041**	0.081	-0.194	0.287	**HML**	**0.080**	0.043	-0.043	0.174
IA	0.033	0.059	-0.101	0.167	**IA**	0.018	0.042	-0.104	0.091
MKT	**0.161**	0.077	-0.039	0.366	**MKT**	**0.086**	0.052	-0.026	0.213
RMW	0.030	0.080	-0.139	0.190	**RMW**	-0.056	0.090	-0.319	0.085
ROE	0.314	0.092	0.018	0.551	**ROE**	0.374	0.113	-0.088	0.546
SMB	**0.021**	0.073	-0.198	0.167	**SMB**	**0.059**	0.049	-0.056	0.191
UMD	**-0.048**	0.056	-0.158	0.138	**UMD**	**-0.124**	0.031	-0.211	-0.062

Notes: The table reports *OOS* RL portfolio weights statistics for each of 9 factors for an investor with monthly rebalancing with short, 12 months (12m), versus long, 120 months (120m) investment horizon. The *OOS* period is from January 2005 to December 2020.

Table 4: *OOS* RL Portfolio Weights - Monthly Rebalancing, Short vs Long Horizon Allocations.

Polk et al. (2022) argue that long term investor should hold lower position in UMD but they do not consider short-selling. Our long term investor does hold lower position in UMD compared to short term investor, it is however more negative, short, position.

Among other factors, the short term investor holds on average twice more of MKT, 16.1%, compared to the long term investor, 8.6%. The long term investor on average holds negative position in RMW, -5.6%, while short term investor keeps small positive weight in RMW, 3%. Out of all factors, MKT has performed the best in terms of cumulative returns over our *OOS* period, Figure 7b, while RMW has a moderate positive performance. Higher weights by short-term investors in these factors can be attributed to trying to benefit from their short-term performance.

Interestingly, both short- and long-term strategies invest similarly into BAB factor, 12% on average. By construction, BAB is market neutral and hence does not have a clear pattern of mean reversion in the long run. Yet, BAB is the second best performing factor in our out-of-sample period, Figure 7b. Giving these properties, both investors choose to invest similarly regardless their investment horizons.

Overall, we find that investment horizon plays a crucial role in a portfolio performance. It has an effect on both, portfolio turnover strategies and factor weights allocations. Consistent with the evidence in the previous literature, factors which have higher expectations of mean reversion in the long run based on their predictability by the state variables, obtain higher weights for long-term horizon investors. This subsequently leads to lower rebalancing needs and portfolio turnover by long term investors, who have longer periods to gradually adjust their portfolio weights to the optimal targets. As a result, when both long- and short-term investors rebalance their portfolios monthly, the portfolio of long-term investors outperforms.

	High Volatility	High Illiquidity	Low Volatility	Low Illiquidity
Return	0.0060	0.0065	0.0084	0.0079
Std.Dev.	0.0126	0.0127	0.0079	0.0080
Sharpe	1.642	1.775	3.652	3.418
6 Factor Alpha	0.0049	0.0049	0.0042	0.0045
6 Factor Alpha T	5.461	4.826	6.450	7.805
6 Factor IR	1.625	1.618	2.419	2.623
9 Factor Alpha	0.0049	0.0048	0.0028	0.0033
9 Factor Alpha T	5.471	4.619	4.203	4.327
9 Factor IR	1.615	1.574	1.807	2.019
Max 1 Period Loss	-0.032	-0.032	-0.008	-0.014

Notes: The table presents the multifactor portfolio out-of-sample, *OOS*, performance for monthly rebalancing and 12 month horizon strategy only, and by market regimes. The market regimes are conditioned on the volatility and illiquidity relizations. High (low) market volatility regime corresponds to the time periods when the values of VIX are above (below) the *OOS* historical median. High (low) market volatility regime corresponds to the time periods when the values of VIX are above (below) the *OOS* historical median. High (low) market illiquidity regime corresponds to the time periods when the values of market-wide aggregate illiquidity are above (below) the *OOS* historical median. The returns, standard deviations and risk-adjusted alphas are monthly, while Sharpe ratios are annualized. 6-factor ALpha and IR (information/appraisal ratio) are estimated from Fama-French 5 factor model plus momentum factor, while 9-factor model includes all 9 factor portfolios in our analysis. IR is computed as Alpha divided by residual standard deviation after regressing portfolio returns on either 6 or 9 factors, and is annualized. The *OOS* period is from January 2005 to December 2020.

Table 5: Out-of-Sample Multifactor Portfolio Performance - Market Regimes

5.3 Market Regimes

The performance across different market regimes is always an "achilles heel" of deep learning models. Avramov et al. (2021) find that the best performances of ML models in the current literature are driven by the periods of high limits to arbitrage, i.e. when the market volatility is high. Besides, there is also the literature which finds that predictability of the main financial indexes, such as market returns of G7 countries, with dividend yields or short interest rates does not exist during business cycle expansions (Henkel et al. (2011)). The authors argue that the predictability is driven by the market downturns or the best identified during business cycle contractions. Relatedly, Avramov et al. (2021) notice that these are the periods where professional investors are the most financially constrained, and hence the value of these signals is hard to arbitrage away.

On the other hand, Barroso and Detzel (2021) argue that the volatility-managed market portfolio of Moreira and Muir (2017) outperforms the market only following months of high sentiment, and it underperforms the market during low-sentiment periods. Therefore, sentiment traders are partially responsible for overall portfolio outperformance. DeMiguel et al. (2021) confirm this observation in-sample, and show even further that the volatility-managed market portfolio underperforms in *OOS* the unmanaged market factor during both low- and high-sentiment periods. The authors argue however that the *OOS* performance of their conditional on volatility multifactor portfolio is better during both regimes than the unconditional passive benchmark.

Table 5 presents portfolio performance statistics for monthly rebalancing and 12 month investment horizon strategy across different market volatility and illiquidity regimes.[13] High (low) market volatility regime corresponds to the months when the values of CBOE-VIX monthly volatiltiy index are above (below) the *OOS*, January 2005 to December 2020, historical median. Given that VIX is often commonly recognized as a market "fear" index, low VIX regimes can also be associated with high investor sentiment periods. Similarly, High (low) market illiquidity regime corresponds to the time periods when the values of market-wide aggregate illiquidity are above (below) the *OOS* historical median. We use Corwin and Schultz (2012) measure of bid-ask spread to aggregate it from the stock to the market wide level as a measure of illiquidity.

Our objective is not to compare conditional vs. unconditional performance but rather to identify performance sensitivity to different market environments. The highest Sharpe ratio is observed for the low volatility regime, 3.65. It decreases almost twice to 1.64 for high volatility regimes. Yet, this Sharpe ratio still outperforms all passive benchmarks or their static combinations as suggested by highly statistically and economically significant Alpha values. These results are also qualitatively similar for high vs. low illiqudity regimes.

Interestingly, the differences in information ratios, IRs, are not that pronounced as those for the Sharpe ratios across different regimes. For example, the 9 factor IR for the low volatility regime is 1.81, which is very close to IR of 1.62 for the high volatility regime. Therefore, the *relative* outperformance of our RL portfolio across all passive benchmarks is relatively similar across different regimes.

We conclude that our portfolio performs consistently better than passive benchmarks or their static combinations across different market conditions.

5.4 *OOS* Portfolio Performance: Lower Re-balancing Frequencies and Longer Horizon Returns

Table 6 presents our multifactor portfolio performance for the *OOS* period from January 2005 to December 2020, where we compare the monthly re-balancing and the base, 12 month horizon (the first column in Table 2) to quarterly, semi-annual and annual re-balancing and 3, 5 and 10 year investment horizons respectively. A critical difference from the previous approach is

[13] The results for longer horizons are qulitatively very similar.

	Monthly	**Quarterly**	**Semi-Annually**	**Annually**
Return	0.0072	0.0188	0.0357	0.0737
Std.Dev.	0.0106	0.0155	0.0166	0.0273
Sharpe	2.344	2.428	3.051	2.702
6-Factor Alpha	0.0053	0.0157	0.0301	0.0600
6-Factor Alpha T	8.254	7.204	9.913	8.770
6-Factor IR	2.133	2.157	2.959	2.923
9-Factor Alpha	0.0051	0.0146	0.0255	0.0645
9-Factor Alpha T	7.636	7.159	8.343	7.842
9-Factor IR	2.086	2.071	2.666	3.241
MaxDD	0.044	0.024	0.011	0.000
Max 1 Period Loss	-0.032	-0.023	-0.011	0.000

Notes: The table presents the multifactor portfolio out-of-sample, *OOS*, performance by holding period/rebalancing frequencies ranging from 1 month to 1 year. The returns, standard deviations and risk-adjusted alphas are scaled to the length of holding periods, while Sharpe ratios are annualized, i.e. the same scale across all frequencies. 6-factor ALpha and IR (information/appraisal ratio) are estimated from Fama-French 5 factor model plus momentum factor, while 9-factor model includes all 9 factor portfolios in our analysis. IR is computed as Alpha divided by residual standard deviation after regressing portfolio returns on either 6 or 9 factors, and is annualized. IR can be viewed as Sharpe ratio when all the factors are hedged out, i.e. the alpha expressed via Sharpe ratio. The *OOS* period is from January 2005 to December 2020.

Table 6: Out-of-Sample Multifactor Portfolio Performance

longer horizon of quarterly, semi-annual and annual holding period returns, respectively. Here, depending on the length of holding period, the returns, standard deviations and risk-adjusted portfolio alphas are either monthly, quarterly, semi-annual or annual respectively.

Similar to the previous results, Table 2, the Sharpe ratios reported in Table 6 are increasing with the holding period. For the monthly rebalancing and 12 month horizon the results are copied from Table 2 for comparison.

The highest Sharpe ratios are obtained for semi-annual, 3.051, and annual, 2.702 holding periods for which the model is trained, in the training sample, to recognize the 5- and 10-year investment horizons, respectively. IR is too increasing with the investment horizon, and for 9-factor model, it monotonically raises from 2.086 with monthly to 3.241 with annual re-balancing.

Interestingly, for annual re-balancing and 10 year horizon where the model is trained on annual holding period returns, the results are very similar to monthly rebalancing and 120 month horizon, where the model is trained on monthly holding period returns (Table 2, the last column).

	Monthly	**Quarterly**	**Semi-Annually**	**Annually**
Avg TC (100% Turnover)	0.0084	0.0084	0.0082	0.0086
Avg TC $\times$ Turnover All	0.0009	0.0010	0.0014	0.0020
Turnover All	0.110	0.119	0.165	0.238
Turnover Long	0.078	0.090	0.123	0.172
Turnover Short	0.032	0.028	0.040	0.065

Notes: The table presents summary statistics for *OOS* RL portfolio turnovers and trading costs. Turnovers are estimated at the frequency of rebalancing ranging from the monthly to annual. Trading costs, TC, are always the average trading costs for the month of rebalancing, i.e. the last month of the holding period. The *OOS* period is from January 2005 to December 2020.

Table 7: *OOS* RL Portfolio Turnover and Trading Costs

One can ask a question of whether this performance is driven by massive rebalancing/turnover, and especially for annual re-balancing, which could simply be impossible to implement in real time (Avramov et al. (2021)). We report summary statistics of turnover and trading costs, TC, in Table 7.

First, if we do have high turnover then the average TC can be prohibitively expensive, and for the monthly rebalancing, the average TC of 84 bps exceed monthly average return of 72 bps. However, our monthly turnover is not 100% - it is 11%, with approximately 8% and 3% on the long and short positions respectively. We therefore pay only 11% of the total trading costs associated with the total portfolio turnover, which results in economically viable and realistic from implementation point of view estimate of 9 bps. Not surprisingly, the highest turnover is observed for the annual rebalancing, 23.8% with the annual trading costs of 20 bps. Note that this is substantially lower than annualized TC of monthly rebalancing of 108 bps ($9bps \times 12$).

Another important question to address is whether our RL portfolio uses unreasonable leverage or extreme portfolio weights (Avramov et al. (2021)) which can drive its superior performance. Table 8 presents summary statistics of portfolio weights and their distributions. These statistics are first computed for each holding frequency in the *OOS* period and then are averaged across these frequencies for the whole sample period. Since we have only 9 assets to invest, the weights here are larger than one would normally observe for individual stocks in the case of bottom-up portfolio construction. Nevertheless, we do not observe any extreme positions which has been reported by Avramov et al. (2021) who analyze other methodologies involving factor investments. Our average position is between 6% to 7.7%. The min and max positions are the smallest in absolute values for the monthly rebalancing, -8.22% and 32.54% respectively. The highest in absolute value short position is -13.94% observed for semi-annual rebalancung, Panel C. These is far away from the extreme leverage positions

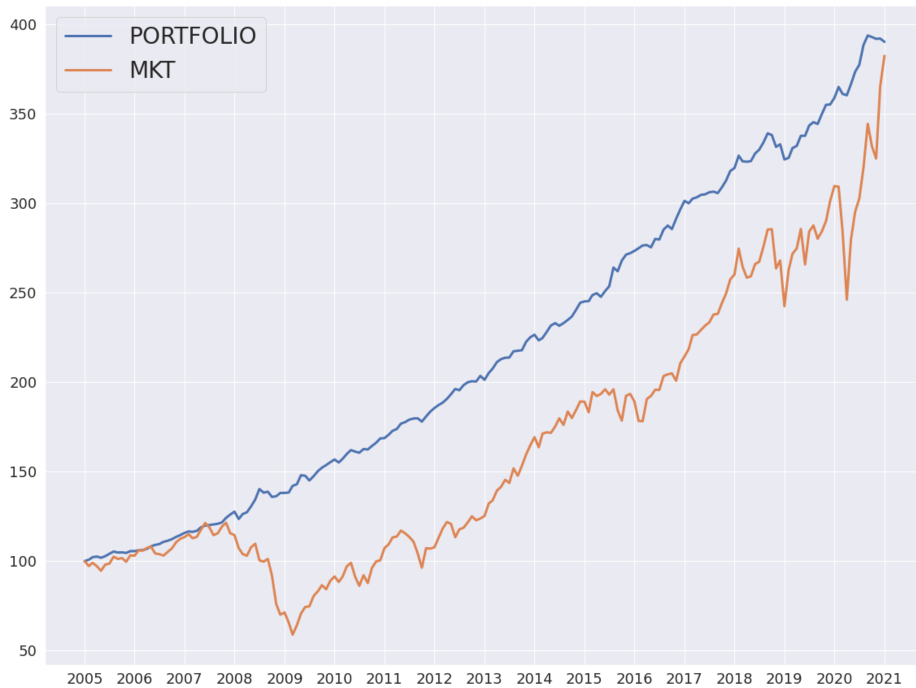
Panel A. Monthly										
Mean	Std.Dev.	Min	5%	10%	25%	Median	75%	90%	95%	Max
7.6730	12.2692	-8.2224	-6.2985	-4.3746	0.0479	4.9023	13.5626	21.5799	27.0593	32.5387
Panel B. Quarterly										
Mean	Std.Dev.	Min	5%	10%	25%	Median	75%	90%	95%	Max
6.8248	13.3560	-12.5471	-8.8400	-5.1328	0.5614	4.5444	11.2623	19.4956	28.0516	36.6076
Panel C. Semi-Annually										
Mean	Std.Dev.	Min	5%	10%	25%	Median	75%	90%	95%	Max
6.1480	13.1031	-13.9432	-10.1865	-6.4297	0.6299	5.0041	9.4812	18.2071	26.6482	35.0893
Panel D. Annually										
Mean	Std.Dev.	Min	5%	10%	25%	Median	75%	90%	95%	Max
6.3886	13.8921	-12.3721	-10.2897	-8.2072	0.9865	4.4470	9.9668	18.8188	28.4710	38.1233

Notes: The table reports *OOS* summary statistics of RL portfolio weights assigned for each holding frequency to 9 factor portfolios comprising the overall RL portfolio level asset allocation. The *OOS* period is from January 2005 to December 2020.

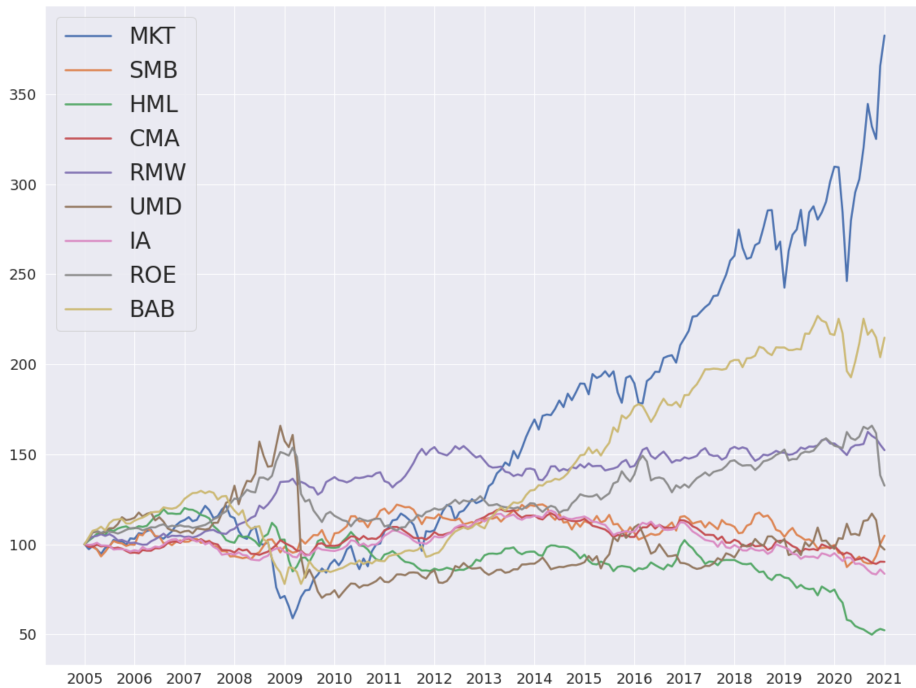
Table 8: Summary Statistics: *OOS* RL Portfolio Weights

exceeding 100% reported by Avramov et al. (2021). The highest long position is observed for the annual rebalancing, 38%, which is still a reasonable extreme give only 9 asset allocation choices to make.

To track the intertemporal performance, Figure 7a plots the monthly *OOS* cumulative return of our RL portfolio for monthly rebalancing and 12 month investment horizon, as the base case specification vis a vis MKT factor, while Figure 7b reports cumulative returns of all 9 factors that we use to construct RL portfolio. Overall, MKT is the best performing factor across all others during our *OOS* period from January 2005 to December 2020. Yet, RL portfolio consistently outperforms MKT, Figure 7a, throughout out the sample, and it has much lower volatility compared to that of MKT. Due to the diversification effect across factors, RL portfolio does not plunge as much during for example 2008-2009 financial crisis, or during March 2020 COVID-19 market dip. Some factors, Figure 7b, are quite stable, e.g. ROE and RMW, and have lower volatility compared to others around crisis events. UMD collapses after the mid-2008 and never comes back to pre-2008 levels. Consistent with popular press observations, value factor, HML, too has been performing poorly during our *OOS* period. BAB collapses around 2008 crisis as well, but then it recovers and finishes the second after MKT. Other factors like SMB, CMA or IA are very close to the straight line in terms of cumulative performance, except 2013-2014 where these factors experience positive gains.



(a) Cumulative returns: RL Portfolio vs MKT factor



(b) Cumulative returns: Factor Portfolios

Figure 7: Cumulative *OOS* Returns: RL Portfolio vs Factors

5.5 Portfolio Policy Rules: *OOS* Performance with Two Assets and No Big Data

What drives the model's performance? It is dynamic factor weights allocations and efficient market/factor timing. To demonstrate this, we set up an empirical experiment the outcomes of which we can anticipate guided by the results of the previous literature.

Here we present the portfolio policy rule using only two assets. MKT is the first natural choice as the most of empirical and theoretical asset pricing literature uses it as the basis asset. We use BAB as the second asset, and as another natural choice. By construction, BAB is market neutral, as it is long leveraged low-beta assets and short high-beta assets. It thus naturally fits for a diversification purpose.

As these assets' characteristics, following Brandt et al. (2009), we use only three: size, book-to-market, and momentum. An important ex ante observation is that we know that these characteristics are important predictors of individual stock returns, and hence useful for individual stocks' bottom-up portfolio construction (Brandt et al. (2009)). They are not as relevant for predicting returns of well diversified portfolios such as MKT and BAB. Therefore, for this particular scenario with the limited predictive information, our model has no advantage compared to a simple static combination of these two factors in a portfolio, or a simple Markowitz approach.

In contrast, there is quite extensive theoretical and empirical literature initiated by Campbell and Shiller (1988) and Fama and French (1988) on predicting market returns, MKT, with dividend-price ratio. Dividend-price, DP, ratio has also been used as a state variable to solve for multi-period portfolio asset allocation problem by Campbell and Viceira (1999) and Campbell et al. (2003), who also favor it for its predictive ability for the market returns. Brennan et al. (1997) and Lynch (2001) use it to condition asset return moments to solve dynamic portfolio programming optimization. We also confirm the previous literature that market returns, MKT, and dividend-price ratios are negatively correlated (Campbell and Viceira (1999), Stambaugh (1999)). BAB and dividend-price ratio are also negatively correlated in our sample.

Guided by the results of the previous literature, we can postulate our main hypotheses. First, in the specification with only three asset characteristics, and the absence of predictive signal, DP, we do not expect any gains to long term portfolio, or any improvement in dynamic combination of the two factors compared to their simple static combination. This specification does not contain market timing signal.

Second, given better predictability of longer horizon returns by dividend-price ratio (Fama and French (1988)), we should expect higher utility gains for longer term portfolios. We should observe higher gains for annual rebalancing in particular, given that the *annual* returns are better predicted by dividend-price ratio (Fama and French (1988), Campbell et al. (2003), Golez and Koudijs (2018)) compared to other higher frequency returns. Since our utility is the utility of wealth of mean variance efficient investor, the utility gains can directly be traced to the Sharpe ratios. Even though we are predicting weights, rather than returns, this stylized return predictability should directly be captured by the portfolio weight allocations (Brandt (1999), and higher Sharpe ratios.

Third, the prior literature provides guidance on the wealth allocation between risky and risk-free assets for different investment horizons, where long term investor allocates more to the risky asset compared to the short term investor (Campbell and Viceira (1999), Campbell et al. (2003)). This is because DP ratio is mean-reverting in the long-run and so are long horizon returns, and the long term investor, thus, benefits the most from this mean-reversion. We cannot speak directly to this literature as we allocate between two risky assets, MKT and BAB. We however can test a portfolio policy rule established by this literature. In particular, given contemporaneous negative correlation between DP and market returns, and positive predictive affect of DP for the future market returns (Brandt (1999), Barberis (2000)), we should expect negative relations between predicted portfolio weights and the actual DP realizations. That is the future asset allocations should reflect these relations, and invest more in risky assets, both MKT and BAB, when DP is expected to be low, while future returns to be high, and visa versa.

Table 9 presents the first sets of results using only three asset specific characteristics. First, Sharpe ratios are decreasing with the horizon, with monthly rebalancing having the highest Sharpe ratio of 0.813. For the reference, for this out-of-sample period, the annualized Sharpe ratios of MKT and BAB are 0.63 and 0.55 respectively. Consistent with diversification principle of Markowitz (1952), bringing these assets in a portfolio improves risk-return trade off only for short-term investor, with monthly rebalancing and 1 year horizon. Moreover, there are no economic gains for what supposed to be an optimal dynamic rebalancing strategy as measured by two-factor, MKT and BAB, risk adjusted portfolio returns, i.e. Alphas. Zero portfolio alphas for all horizons indicate that any passive, static combination of these two factors provide similar performance. Indeed, judging by extremely low portfolio turnovers, e.g. 2.5% for monthly rebalancing, one can conclude that the model almost does not trade at all. Portfolio weights statistics reported in the lower panel of Table 9 have extremely low standard deviations confirming the turnover evidence that the model does not actively time the market. This is expected as there are no state variables, i.e. dividend-price ratio, which would allow for the market predictability and timing. The model allocates more weights to BAB rather than to MKT, and the gap between the weights gets wider for the longest

	Monthly	Quarterly	Semi-Annually	Annually
Return	0.0059	0.0121	0.0365	0.0625
Std.Dev.	0.0253	0.0532	0.0820	0.1354
Sharpe	0.813	0.454	0.630	0.462
2 Factor Alpha	0.0000	0.0000	-0.0006	-0.0070
2 Factor Alpha T	0.044	0.006	-0.314	-0.919
2 Factor IR	0.014	0.001	-0.067	-0.158
MaxDD	0.544	0.503	0.510	0.470
Max 1 Period Loss	-0.119	-0.187	-0.299	-0.360
Turnover	0.025	0.063	0.053	0.057
Portfolio weights				
			BAB	
mean	0.527	0.529	0.556	0.575
Std.Dev.	0.068	0.155	0.075	0.088
min	0.368	-0.353	0.470	0.507
max	0.963	0.856	0.803	0.791
			MKT	
mean	0.465	0.386	0.444	0.399
Std.Dev.	0.108	0.262	0.075	0.177
min	-0.405	-0.880	0.197	-0.214
max	0.632	0.511	0.530	0.493

Notes: The table presents a two factors, MKT and BAB, RL portfolio out-of-sample, *OOS*, performance by holding period/rebalancing frequencies ranging from 1 month to 1 year. The factor weights allocation is conditioned on three asset specific characteristics: size, book-to-market and momentum. The returns, standard deviations and risk-adjusted alphas are scaled to the length of holding periods, while Sharpe ratios are annualized, i.e. the same scale across all frequencies. 2-factor Alpha and IR (information/appraisal ratio) are estimated from MKT and BAB 2 factor model. IR is computed as Alpha divided by residual standard deviation after regressing portfolio returns on 2 factors, and is annualized. IR can be viewed as Sharpe ratio when all the factors are hedged out, i.e. the alpha expressed via Sharpe ratio. The *OOS* period is from January 2005 to December 2020.

Table 9: *OOS* Two Factor RL portfolio performance with only three asset specific characteristics

horizon with the annual rebalancing. Here the average weight of BAB is 0.58, while it is 0.40 for MKT.

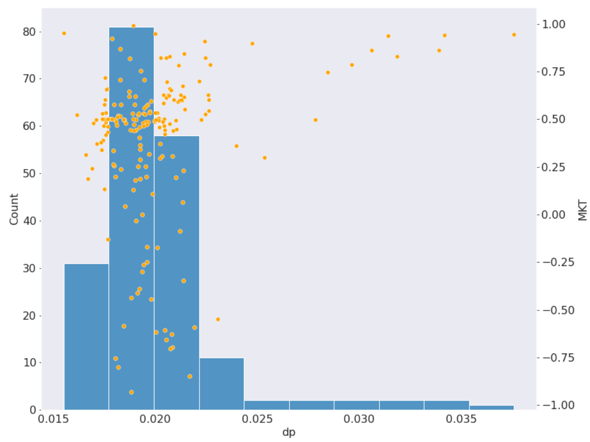
The results in Table 10 present a completely different picture. The estimation here includes one state variable, the dividend-price ratio. As return predictability improves with the horizon (Fama and French (1988), Brandt (1999), Golez and Koudijs (2018)), so do, almost monotonically, the Sharpe ratios. Here, the Sharpe ratio for the monthly holding period is

	Monthly	Quarterly	Semi-Annually	Annually
Return	0.0050	0.0116	0.0346	0.1120
Std.Dev.	0.0302	0.0585	0.0712	0.0965
Sharpe	0.578	0.397	0.688	1.161
2 Factor Alpha	0.0017	0.0018	0.0337	0.1260
2 Factor Alpha T	0.610	0.218	2.790	4.799
2 Factor IR	0.229	0.070	0.665	1.214
MaxDD	0.636	0.588	0.166	0.033
Max 1 Period Loss	-0.150	-0.188	-0.108	-0.030
Turnover	0.143	0.226	0.263	0.333
Portfolio weights				
			BAB	
mean	0.376	0.396	0.400	0.398
Std.Dev.	0.367	0.384	0.416	0.479
min	-0.937	-0.670	-0.699	-0.826
max	0.999	0.992	0.827	0.751
			MKT	
mean	0.387	0.317	0.219	0.352
Std.Dev.	0.399	0.408	0.414	0.224
min	-0.929	-0.913	-0.832	-0.201
max	0.990	0.858	0.644	0.534

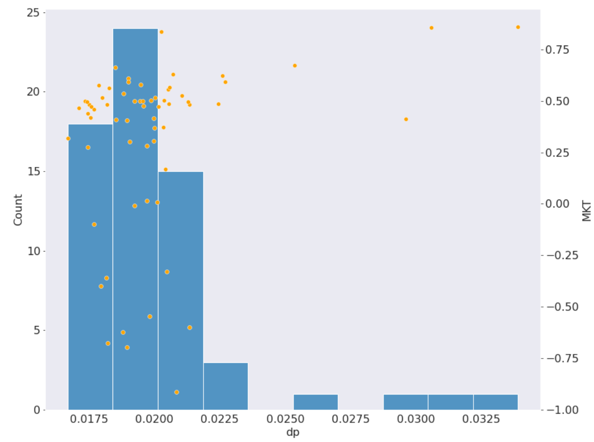
Notes: The table presents a two factors, MKT and BAB, RL portfolio out-of-sample, *OOS*, performance by holding period/rebalancing frequencies ranging from 1 month to 1 year. The factor weights allocation is conditioned on three asset specific characteristics: size, book-to-market and momentum, as well as one state variable, dividend-price ration, DP. The returns, standard deviations and risk-adjusted alphas are scaled to the length of holding periods, while Sharpe ratios are annualized, i.e. the same scale across all frequencies. 2-factor Alpha and IR (information/appraisal ratio) are estimated from MKT and BAB 2 factor model. IR is computed as Alpha divided by residual standard deviation after regressing portfolio returns on 2 factors, and is annualized. IR can be viewed as Sharpe ratio when all the factors are hedged out, i.e. the alpha expressed via Sharpe ratio. The *OOS* period is from January 2005 to December 2020.

Table 10: *OOS* Two Factor RL portfolio performance with three asset specific characteristics, and one state variable.

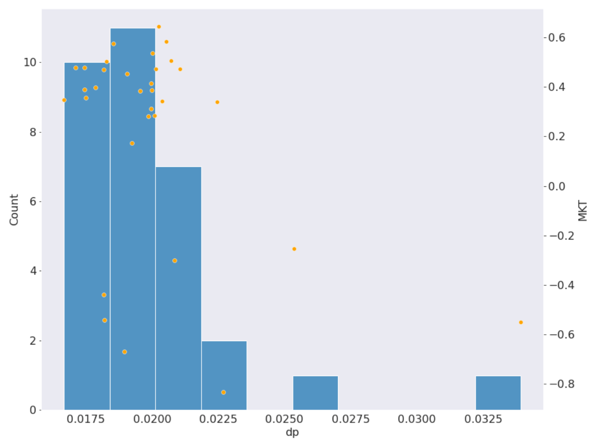
0.578, and it jumps to 1.16 for the annual rebalancing and 10 year horizon investor. Moreover, we also observe significant positive Alphas for the semi-annual and annual rebalancing. Therefore, dividend-price ratio predictability of the market returns over longer horizons allows for dynamic market timing. This in turn provides superior performance to a myopic, static portfolio. The turnovers here are substantially higher as the model trades more given the availability of predictive signals. The standard deviations of portfolio weights are as high as



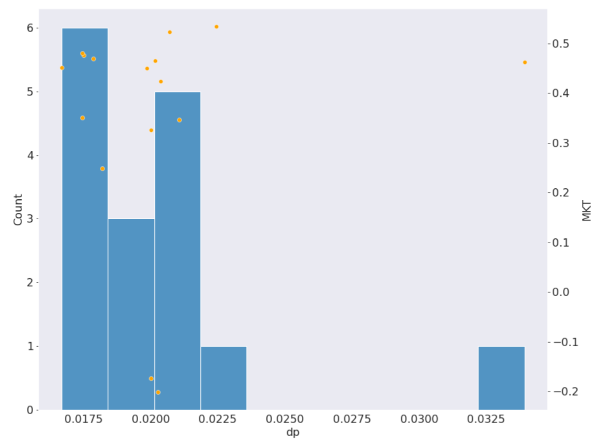
(a) Monthly rebalancing



(b) Quarterly rebalancing



(c) Semiannual rebalancing



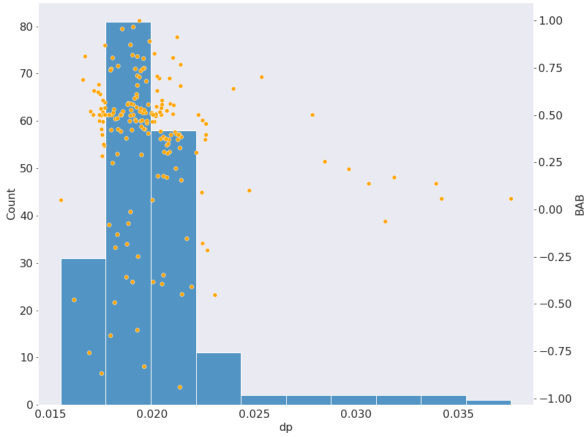
(d) Annual rebalancing

Figure 8: *OOS* Dynamic MKT Predicted Weights vs. Histogram of Dividend-Price ratio Realizations

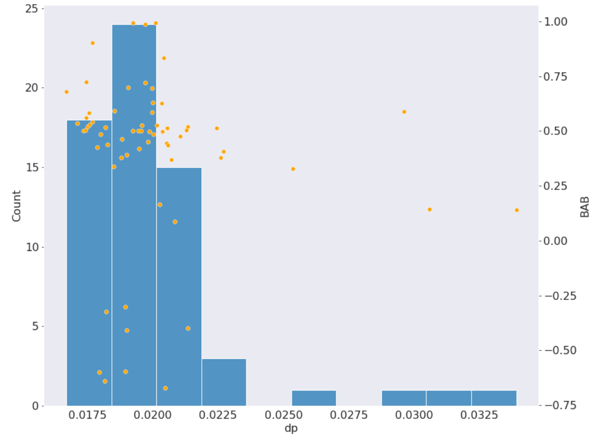
their means. For longer horizons the model continues outweigh BAB compared to MKT.

To summarize, without predictive signals, the model does not outperform a myopic portfolio strategy, Table 9, does not trade or rebalance a lot, and fails to time the market. Once there is a predictor state variable, Table 10, the model starts aggressive rebalancing to time the market and generate significant abnormal returns, alphas, over longer horizons where the return predictability is better.

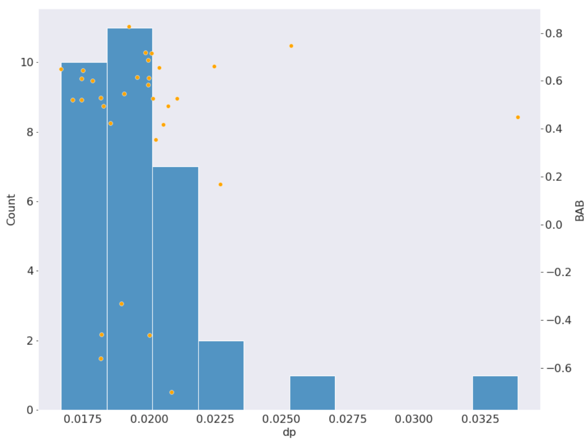
To test our final, third hypothesis about relations between predicted asset allocations and DP realizations, Figure 8 and Figure 9 plot MKT and BAB portfolio weights respectively as a function of DP histogram realizations. Here, for every rebalancing frequency, the weights are the *OOS* predicted weights for the next month, quarter, half-year or year, while DP is the monthly DP value realized in the end of each respective frequency.



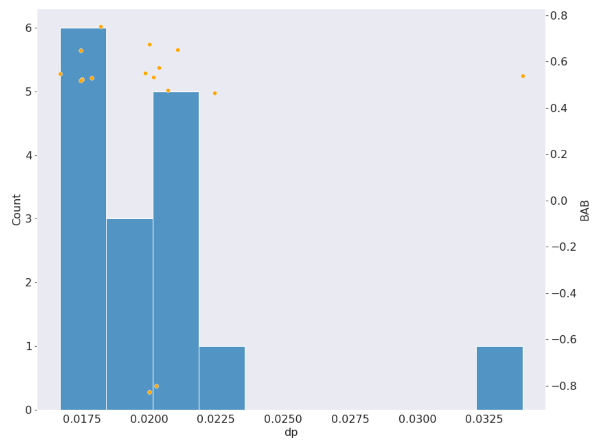
(a) Monthly rebalancing



(b) Quarterly rebalancing



(c) Semiannual rebalancing



(d) Annual rebalancing

Figure 9: *OOS* Dynamic BAB Predicted Weights vs. Histogram of Dividend-Price ratio Realizations

For example, Figure 8a presents monthly MKT predicted weights, orange dots, for *OOS*, 01/2005 -12/2020, period versus the histogram of DP realizations for these months, blue bar-chart. As one can easily notice, the majority of weights (dots) is clustered at the lower percentiles of DP frequency realizations. This relation is the best characterized by the density of cluster rather than linearity argued by the previous literature (Brandt (1999)).

Similar relations are observed for the monthly BAB predicted weights and DP realizations in Figure 9a. Yet, one can fit here a negatively sloped straight line between DP and BAB-weights relations (Brandt (1999)), where the most of weights still clustered at the lower percentiles of DP frequency realizations. Other sub-figures in Figure 8 and Figure 9 exhibit similar relations between the asset allocation weights and DP for lower rebalancing frequencies.

Overall we conclude that portfolio rule of our model follows the one established by the previous literature: invest more in the risky assets when the realization of the state variable, DP, is low, and visa versa (Brandt (1999), Campbell and Viceira (1999), Campbell et al. (2003)). The difference from the previous literature is that this rule is neither liner not quadratic, but rather better characterized by clustering and its density or the polynomial structure which machine learning algorithms identify in the data driven way.

The on-line Appendix B presents factor timing use-cases around COVID-19 March 2020 market crash, and 2008-2009 financial crisis episode. In both cases the model takes a short position a few months before one of the factors crashes, which further supports its market timing ability.

6 Economic Determinants of Performance

6.1 Variable Importance

The optimization leverages a lot of factor-specific features, macro-variables, and interactions between them. Which factor-specific characteristics or macro-economic variables matter the most? We use the *saliency map* technique from Simonyan et al. (2013) that attributes feature importance for a single prediction by taking the absolute value of the partial derivative of the output with respect to the input features. The absolute value of the gradient points to input features that can be perturbed the least in order for the target output to change the most: the higher it is for a variable, the more it plays a role in the forecast. By producing the saliency map for all assets over the whole training sample, we can rank what features are the most important based on the overall importance-score $S(\mathbf{x})$ computed for each feature.

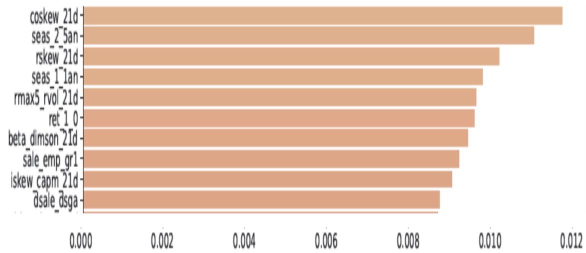
More specifically, given the i-th $\mathbf{x}_{t,i} \in \mathbb{R}^D$ observed at time t, the sensitivity (the importance) $S(\mathbf{x}_{i,t}) \in \mathbb{R}^D$ of the D's features can be estimated with eq.11.

$$S(\mathbf{x}_{t,i}) = \left| \frac{\partial f_\theta(\mathbf{x}_{t,i})}{\partial \mathbf{x}_{t,i}} \right| \quad (11) \qquad S(\mathbf{x}) = \texttt{pool}\left(\sum_{t=1}^{T} \sum_{i=1}^{N_T} S(\mathbf{x}_{t,i}) \right) \quad (12)$$

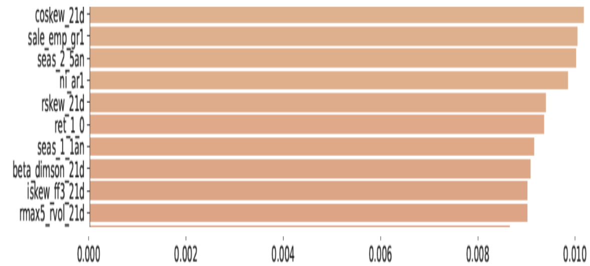
where t is an OOS period (month, quarter, half-year, year), T is the number of OOS periods, and N_T is the number of assets at time t.

The overall ranking can be computed from eq. 12 where `pool` is a pooling function. We consider the median to pool all importance-scores assigned to the feature across all training samples.

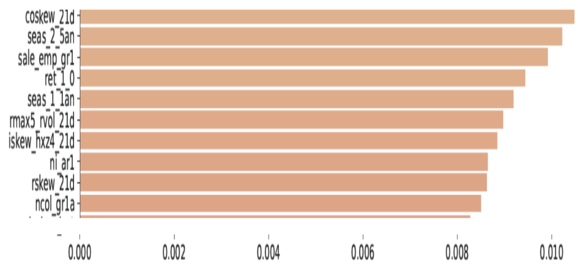
Figure 10 presents top 10 factor-specific variable weights' importance for monthly rebalancing and 12, 36, 60 and 120 month investment horizons in panels 10a, 10b, 10c, and 10d respectively. On average, the rankings between horizons are similar except a few noticeable exceptions. For 12 month horizon, the top 5 variables are completely attributed to the returns



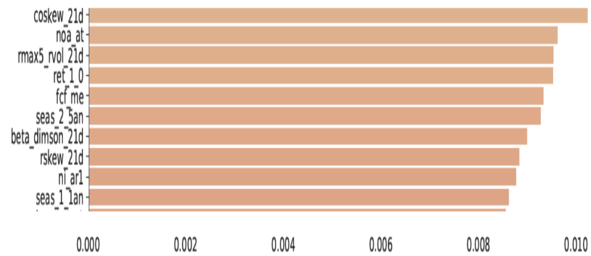
(a) Top 10 predictors for 12 month horizon



(b) Top 10 predictors for 36 month horizon



(c) Top 10 predictors for 60 month horizon



(d) Top 10 predictors for 120 month horizon

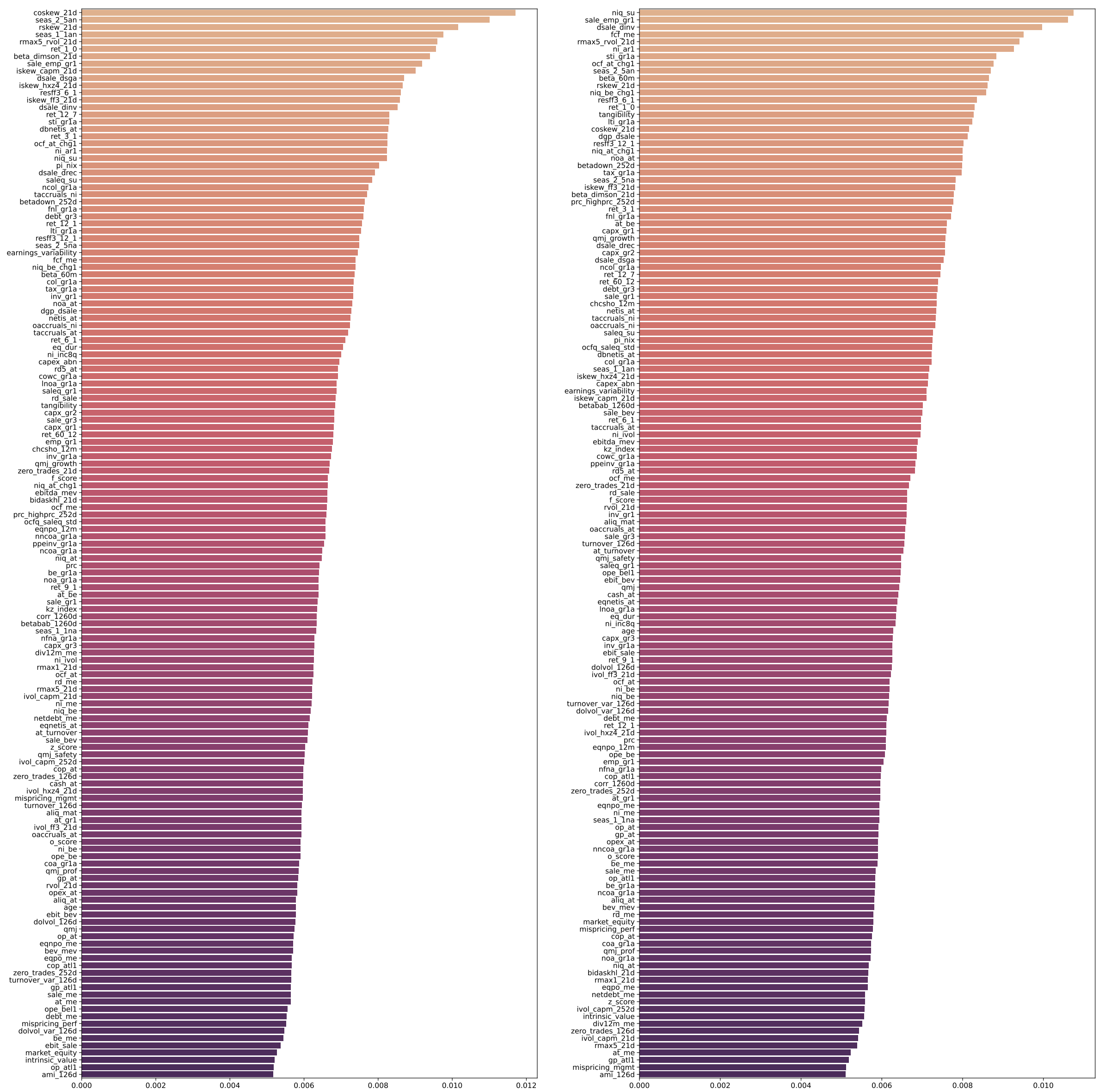
Figure 10: Top 10 predictors for monthly rebalancing and different investment horizons

moments. The firm fundamental variables like sale_emp_gr1 (labour force efficiency) or dsale_dgga (Change sales minus change SG&A) appear in the bottom 5. In contrast, noa_at (net operating assets) and fcf_me (free cash flow-to-price ratio) are in the top 5 for 120 month horizon.

The bigger contrast between variable importance is observed for different re-balancing frequencies. Figures 11a, 11b, 12a, and 12b present the variable weight importance for monthly, quarterly, semi-annual and annual rebalancing respectively. The variable importance is measured by the ranking score results, where all ranks are normalized to add up to 1.

The most important observation is that the top importance variables for the monthly rebalancing and the base case, 12 month horizon, are much different from lower frequency rebalancing. For example co-skewnes, skewness, betas, return volatility, mutiple specifications of idiosyncratic skewness or reversals are clearly dominating the top predictors for monthly rebalancing strategy. In contrast, earnings surprises, 1 year sales growth, changes in sales to inventory ratios, free cash flow to price, earning persistance, net-operating assets, tax expense surprise, Ebitda-to-market enterprise value are the top predictors for lower rebalancing frequencies which take into account 3 to 10 year investment horizon period considerations.

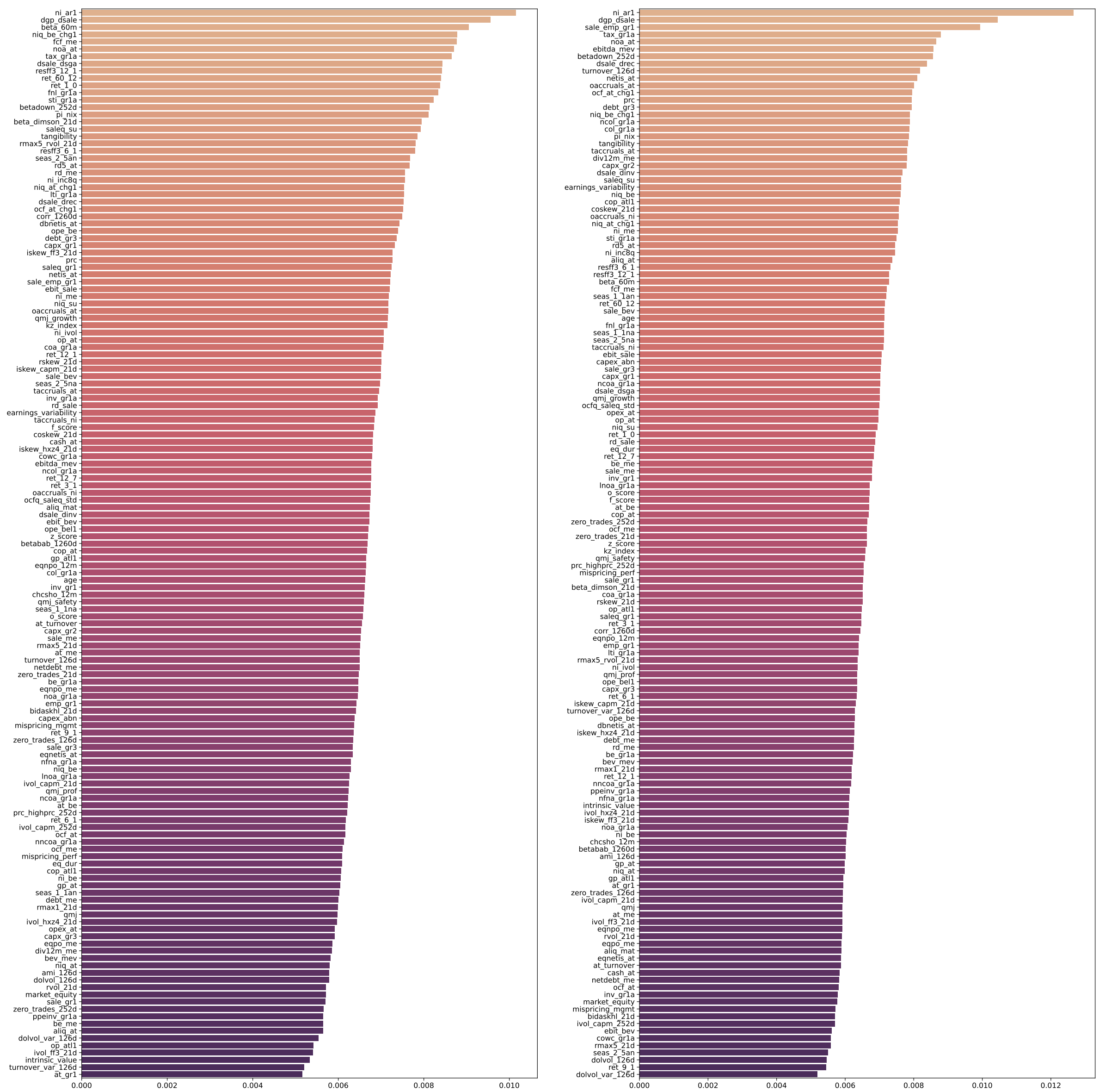
For example the biggest contrast with the monthly rebalancing, Figure 11a, is observed for the annual rebalancing, Figure 12b. Here, none of top 20 important features for annual



(a) Monthly Rebalancing (b) Quarterly Rebalancing

Figure 11: Stock Variable Importance: Monthly and Quaterly

rebalancing overlaps with top 20 features for monthly rebalancing. For annual rebalancing, the first important characteristic is Earnings Persistance, *ni_ ar1*, followed by Change gross margin minus change sales, *dgp_ dsale*, and then by Labor force efficiency, *sale_ emp_ gr1*. The last two variables in the top 5 are Tax expense surprise, *tax_ gr1a*, and Net operating assets, *noa_ at*. For comparison, the top 5 ranks in order of importance for monthly rebalancing are Coskewness, *coskew_ 21d* (Harvey and Siddique (2000)), Years 2-5 lagged returns, annual, *seas_ 2_ 5an* (Heston and Sadka (2008)), Total skewness, *rskew_ 21d* (Bali et al. (2016)),

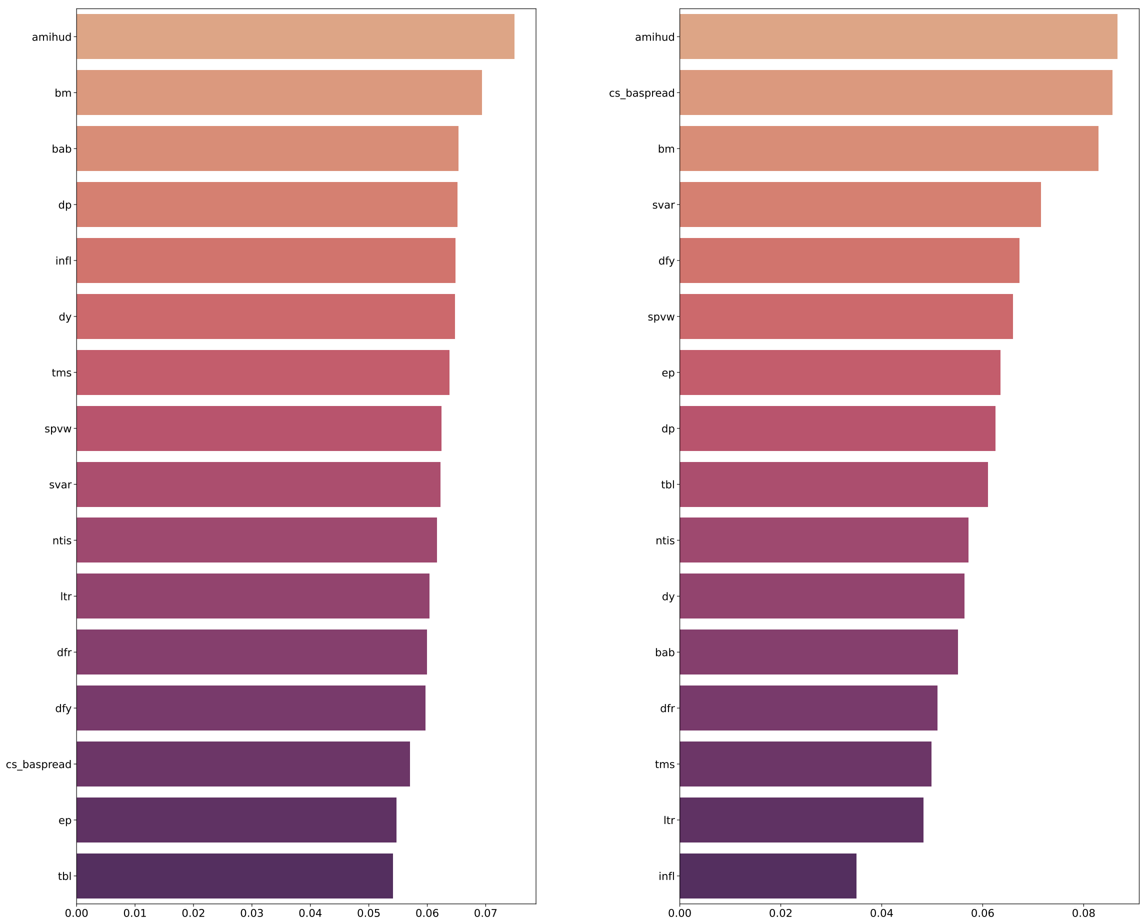


(a) Semi-Annual Rebalancing (b) Annual Rebalancing

Figure 12: Stock Variable Importance: Semi-Annual and Annual

Year 1-lagged return, annual, *seas_1_1an* (Heston and Sadka (2008)), and Highest 5 days of return in a month, *rmax5_21d* (Bali et al. (2017)). These features are not even in the top 40 for annual rebalancing.

This further shows how longer investment horizon influences economic indicators selections. It is purely statistical variables which use return data for the short-term rebalancing, and it is mostly future cash flows indicators, balance sheet, and accounting forward looking profitability for long-term holding periods.



(a) Monthly Rebalancing (b) Annual Rebalancing

Figure 13: Macro-economic Variable Importance

Macro-Economic Variables Macro-economic variables that we use are less important after asset specific features as our factor portfolios can themselves be viewed as macro-factors given the market capitalization and the size of market segments they capture. Moreover, some factors, given their nature, enter into portfolio optimization as assets, and as macro-features. For example BAB is the part of 9 factor portfolio and it is also funding liquidity macro-feature. Similarly, MKT (HML) is one of 9 factors, and it is highly correlated with S&P500 (Book-to-Market ratio), which is a macro-feature as well.

Nevertheless, the 9 factors do not span all macro-economic variables we use, and it is of interest to understand what macro-signals are of the firs order of importance. To contrast short-term vs. long term investment horizons, Figure 13 reports the macro-variable rankings for the monthly, Figure 13a, and annual, Figure 13b, rebalancing only.

Aggregate Amihud Illiquidity measure is always identified as the top macro-feature for both, and the related to it Corwin and Schultz (2012) *cs_spread* illiquidity measure is the second important variable for the annual rebalancing. This result is worth further discussion as our objective function, eq. 9, incorporates asset-specific illiquidity penalty, and hence our optimization is predicting each factor's illiquidity. Asset-specific illiquidity, however, co-moves with the market-wide illiquidity (Chordia et al. (2000)). Therefore, the first order importance of the market wide illquidity, after asset specific features, is reasonably justifiable.

Both strategies highly value market-wide book-to-market ratio, bm. This is too not surprising as even the monthly rebalancing takes into account a one year holding period horizon where the value and associated with it lower cash flow volatility are important factors.

A significant difference appears for BAB ranking, where the monthly rebalancing strategy ranks it in the third place, and the annual - in the 12th place. Since monthly rebalancing evolves more frequent portfolio turnovers, both market-wide illiquidity and closely associated with it funding illiqidity (Brunnermeier et al. (2009)) are of more importance than for the annual rebalancing.

Interestingly, the annual rebalancing, after market illiquidity and book-to-market, prioritizes stock market variance, *svar*. This result is consistent with Moreira and Muir (2019) conclusions that the long-term investor should time the market-wide volatility in order to improve her portfolio performance. This proposition is supported by the data and our RL portfolio modelling approach.

Another interesting observation, in the light of a record high inflation as of the writing of this paper, the monthly rebalancing ranks inflation highly, the 5th in the ranking, while the annual rebalancing with 10 year holding period investment horizon, ranks it the last in the ordering among all 16 macro-features. This again makes perfect economic sense as inflation expectations are more imminent for shorter investment horizons, and the inflation risk today has less impact on portfolio performance in 10 years from today.

Overall, similar to asset specific features, the length of investment horizon makes a significant difference for the importance of macro-economic signals.

6.2 Turnover and Liquidity Constraints

In this section we demonstrate another source of performance - long-horizon portfolio re-balancing and turnover strategies, which are also associated with more strategic factor timing. Our objective function, eq. 9, explicitly incorporates turnover and trading costs penalty. It allows identifying optimal trading strategies gradually towards, in spirit of Gârleanu and Pedersen (2013), "target" portfolio. That is, instead of aggressively trading towards the "target" in one period, the strategy is trained to "spread" the trades over multiple periods. The longer the investment horizon, the more "patient" and strategic trading is expected to be.

It is relatively straightforward to test whether "patience matters" in our settings. Once we remove turnover and illiquidity penalty in the objective function, i.e. the last term of eq. 9, then both approaches, long- vs short-term horizon, become equally impatient. That is regardless of the investment horizon, both will only try to maximize returns while minimizing the portfolio volatility exposure. Thus the objective function reduces to the conditional Markowitz type optimization.

Table 11 report results for monthly re-balancing and various holding period horizons. They are to be compared to those in Tables 2 and 3. In contrast to Table 2, the Sharpe ratios are now monotonically decreasing with an investment horizon. Now, the 12 month horizon strategy becomes dominant, both in terms of performance, portfolio risk measures, and turnovers. Compared to Table 2, the Sharpe ratio for the monthly rebalancing is higher, 2.504 vs. 2.344. This is expected as the liquidity constraints are removed. The long 120 month horizon strategy is affected the most. Its Sharpe ratio drops from 2.74, Table 2 to 1.796. Its overal turnover increases from 6.1% per month, Table 3, to 20.4% per month, a 234% increase. That is without spreading the trades and factor timing over longer horizons strategically, the advantage of long-horizon investment consideration completely disappears.

Table 12 reports the *OOS* similar portfolio performance results for lower rebalancing frequancies, while we still keep monthly rebalancing and 12 month horizon as the base. Given that optimal asset allocation is no longer disciplined to be achieved over multiple annual period horizons, the Sharpe ratio for annual rebalancing drops from 2.702, Table 6, to 2.491, and is similar to the monthly rebalancing of short term investors. This further shows that long term liquidity and turnover consideration, and their respective constraints in portfolio optimization, eq. 9, are crucial for long term portfolio asset management superior performances. This also demonstrates how Markowitz type approach to portfolio management favors short-term investment strategies, and offer no solutions for long-term investors.

The absolute values of turnover do not look explosively high. This is due to multiple characteristics of each factor that we use and which implicitly reduce turnovers (DeMiguel

	12m	**36m**	**60m**	**120m**
Return	0.0053	0.0050	0.0050	0.0046
Std.Dev.	0.0073	0.0076	0.0076	0.0089
Sharpe	2.504	2.287	2.277	1.796
6 Factor Alpha	0.0052	0.0051	0.0053	0.0050
6 Factor Alpha T	9.844	9.525	8.762	7.774
6 Factor IR	2.450	2.369	2.441	1.970
9 Factor Alpha	0.0050	0.0048	0.0052	0.0047
9 Factor Alpha T	8.985	8.584	8.494	7.158
9 Factor IR	2.418	2.338	2.449	1.913
MaxDD	0.037	0.041	0.036	0.083
Max 1 Period Loss	-0.014	-0.017	-0.017	-0.032
Turnover All	0.155	0.164	0.179	0.204
Turnover Long	0.093	0.095	0.101	0.113
Turnover Short	0.061	0.069	0.078	0.091
Turnover All Ch.%	41%	110%	175%	234%
Turnover Long Ch.%	19%	64%	115%	169%
Turnover Short Ch.%	91%	245%	359%	406%

Notes: The table presents the multifactor portfolio out-of-sample, *OOS*, performance for monthly re-balancing, various investment horizons ranging from 12 to 120 months, and no turnover and liquidity penalty in the objective function, eq 9. 6-factor ALpha and IR (information/appraisal ratio) are estimated from Fama-French 5 factor model plus momentum factor, while 9-factor model includes all 9 factor portfolios in our analysis. IR is computed as Alpha divided by residual standard deviation after regressing portfolio returns on either 6 or 9 factors, and is annualized. *Turnover Ch.%* presents percentage turnover increase compared to the base case turnovers with turnover and liquidity constraints reported in Table 2. The *OOS* period is from January 2005 to December 2020.

Table 11: Out-of-Sample Multifactor Portfolio Performance - Monthly Rebalancing,No Turnover and Liquidity Constraints

et al. (2020), DeMiguel et al. (2021)). The percentage increases in turnover, compared to the base case, do, however, look high. For example the highest increase in portfolio turnover is observed for quarterly re-balancing, *All*, 78%, with an increase in turnover on the short positions of 200%. Overall, the short positions turnovers experience the highest increase. For example, for semi-annual rebalancing, the overall increase in turnover is 49%. The turnover of the Long positions increases by 14% while the turnover of the short position by 163% compared to the base case in Table 7. Given that our asset classes are factor portfolios, which are composed of very large market segments, any small increase in turnover can results in

	Monthly	**Quarterly**	**Semi-Annually**	**Annually**
Return	0.0053	0.0169	0.0269	0.0627
Std.Dev.	0.0073	0.0144	0.0184	0.0252
Sharpe	2.504	2.340	2.067	2.491
6 Factor Alpha	0.0052	0.0174	0.0226	0.0459
6 Factor Alpha T	9.844	7.463	8.115	4.555
6 Factor IR	2.450	2.316	1.783	1.744
9 Factor Alpha	0.0050	0.0173	0.0203	0.0475
9 Factor Alpha T	8.985	7.295	7.675	3.790
9 Factor IR	2.418	2.309	1.595	1.817
MaxDD	0.037	0.013	0.015	0.000
Max 1 Period Loss	-0.014	-0.013	-0.015	0.000
Turnover All	0.155	0.212	0.246	0.319
Turnover Long	0.093	0.127	0.140	0.173
Turnover Short	0.061	0.084	0.105	0.138
Turnover All Ch.%	41%	78%	49%	34%
Turnover Long Ch.%	19%	41%	14%	1%
Turnover Short Ch.%	91%	200%	163%	112%

Notes: The table presents the multifactor portfolio out-of-sample, *OOS*, performance by holding period/rebalancing frequencies ranging from 1 month to 1 year. The returns, standard deviations and risk-adjusted alphas are scaled to the length of holding periods, while Sharpe ratios are annualized, i.e. the same scale across all frequencies. 6-factor ALpha and IR (information/appraisal ratio) are estimated from Fama-French 5 factor model plus momentum factor, while 9-factor model includes all 9 factor portfolios in our analysis. IR is computed as Alpha divided by residual standard deviation after regressing portfolio returns on either 6 or 9 factors, and is annualized. Turnovers are estimated at the frequency of rebalancing ranging from the monthly to annual. *Turnover Ch.%* presents percentage turnover increase compared to the base case turnovers with turnover and liquidity constraints reported in Table 7. The *OOS* period is from January 2005 to December 2020.

Table 12: Out-of-Sample Multifactor Portfolio Performance - No Turnover and Liquidity Constraints

very large market wide price impacts. These magnitudes of turnover increase, and especially on the short positions where the limits to arbitrage are particularly high (Avramov et al. (2021)), can be viewed as inadmissible from LTA point of view.

Overall our results support Gârleanu and Pedersen (2013) argument about dynamic portfolio with trading costs considerations. Moreover we also observe that dynamic trading over multiple horizons not only allows to gradually reach "the target portfolio" but also highlights the benefits of longer investment horizon considerations.

We further conclude that turnover and liquidity constraints are not only important modelling choices for the long term investment horizon investors to outperform their short term counterparts, but they are also necessary choices to stand against LTA criticism for both short-term and long-term investment horizon considerations.

7 Conclusions

In this paper we provide a framework for the long-term investors to form and manage their portfolios. Our framework uses Reinforcement Learning, RL, approach, relies on most recent AI tools and explicitly incorporates long, up to 10 years, investment horizon, and big data.

Our portfolio exhibits superior out-of-sample performance with low turnover and trading costs, and avoids extreme individual asset positions or high leverage.

Our results have practical and policy implications. For example pension fund or educational endowments portfolios underperformed passive investment by approximately 1% or 1.6% a year, respectively, for the ten years ending June 30, 2018, and this underperformance is expected to persist in the years ahead (Cochrane (2022)). We suggest that long-term investor underperformance does not need to become the new norm, and long-term investors such as pension funds, can outperform their passive benchmarks and even more active, short-term-oriented peers.

Importantly, we also suggest that any portfolio of a risk averse investor should have long-term investment horizon considerations.

Finally, we can also speak to the recent technological advancement and applications of AI in investment management. After the public release of ChatGPT, there has been an increased practitioners and academics' interest of what this type of technology can do in finance. We are the first to apply the back-bone technology supporting ChatGPT for the most basic problem in investments - a multi-factor and multi-horizon portfolio management. Our results support the premise of even wider future applications of this technology in financial services industry.

References

Aït-sahali, Y. and Brandt, M. W. (2001). Variable selection for portfolio choice. The Journal of Finance, 56(4):1297–1351.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. Journal of financial markets, 5(1):31–56.

Avramov, D., Cheng, S., and Metzker, L. (2021). Machine learning versus economic restrictions: Evidence from stock return predictability. Available at SSRN 3450322.

Balduzzi, P. and Lynch, A. W. (1999). Transaction costs and predictability: Some utility cost calculations. Journal of Financial Economics, 52(1):47–78.

Bali, T. G., Brown, S. J., Murray, S., and Tang, Y. (2017). A lottery-demand-based explanation of the beta anomaly. Journal of Financial and Quantitative Analysis, 52(6):2369–2397.

Bali, T. G., Engle, R. F., and Murray, S. (2016). Empirical asset pricing: The cross section of stock returns. John Wiley & Sons.

Barberis, N. (2000). Investing for the long run when returns are predictable. The Journal of Finance, 55(1):225–264.

Barroso, P. and Detzel, A. (2021). Do limits to arbitrage explain the benefits of volatility-managed portfolios? Journal of Financial Economics, 140(3):744–767.

Bianchi, D., Büchner, M., and Tamoni, A. (2021). Bond risk premiums with machine learning. The Review of Financial Studies, 34(2):1046–1089.

Brandt, M. W. (1999). Estimating portfolio and consumption choice: A conditional euler equations approach. The Journal of Finance, 54(5):1609–1645.

Brandt, M. W. and Santa-Clara, P. (2006). Dynamic portfolio selection by augmenting the asset space. The journal of Finance, 61(5):2187–2217.

Brandt, M. W., Santa-Clara, P., and Valkanov, R. (2009). Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. The Review of Financial Studies, 22(9):3411–3447.

Brennan, M. J., Schwartz, E. S., and Lagnado, R. (1997). Strategic asset allocation. Journal of Economic dynamics and Control, 21(8-9):1377–1403.

Brunnermeier, M., Crockett, A., Goodhart, C. A., Persaud, A., Shin, H. S., et al. (2009). The fundamental principles of financial regulation, volume 11. ICMB, Internat. Center for Monetary and Banking Studies.

Campbell, J. Y., Chan, Y. L., and Viceira, L. M. (2003). A multivariate model of strategic asset allocation. Journal of financial economics, 67(1):41–80.

Campbell, J. Y. and Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. The Review of Financial Studies, 1(3):195–228.

Campbell, J. Y. and Viceira, L. M. (1999). Consumption and portfolio decisions when expected returns are time varying. The Quarterly Journal of Economics, 114(2):433–495.

Carhart, M. M. (1997). On persistence in mutual fund performance. The Journal of finance, 52(1):57–82.

Chatigny, P., Goyenko, R., and Zhang, C. (2021). Asset pricing with attention guided deep learning. Available at SSRN.

Chen, L., Pelger, M., and Zhu, J. (2020). Deep learning in asset pricing. Available at SSRN 3350138.

Chordia, T., Roll, R., and Subrahmanyam, A. (2000). Commonality in liquidity. Journal of financial economics, 56(1):3–28.

Cochrane, J. H. (2022). Portfolios for long-term investors. Review of Finance, 26(1):1–42.

Cong, L. W., Tang, K., Wang, J., and Zhang, Y. (2021). Alphaportfolio: Direct construction through deep reinforcement learning and interpretable ai. Available at SSRN 3554486.

Corwin, S. A. and Schultz, P. (2012). A simple way to estimate bid-ask spreads from daily high and low prices. The Journal of Finance, 67(2):719–760.

DeMiguel, V., Martin-Utrera, A., Nogales, F. J., and Uppal, R. (2020). A transaction-cost perspective on the multitude of firm characteristics. The Review of Financial Studies, 33(5):2180–2222.

DeMiguel, V., Martin-Utrera, A., and Uppal, R. (2021). A multifactor perspective on volatility-managed portfolios. Available at SSRN.

Fama, E. F. and French, K. R. (1988). Dividend yields and expected stock returns. Journal of financial economics, 22(1):3–25.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. Journal of financial economics, 116(1):1–22.

Frazzini, A. and Pedersen, L. H. (2014). Betting against beta. Journal of Financial Economics, 111(1):1–25.

Gârleanu, N. and Pedersen, L. H. (2013). Dynamic trading with predictable returns and transaction costs. The Journal of Finance, 68(6):2309–2340.

Golez, B. and Koudijs, P. (2018). Four centuries of return predictability. Journal of Financial Economics, 127(2):248–263.

Gu, S., Kelly, B., and Xiu, D. (2020). Empirical asset pricing via machine learning. The Review of Financial Studies, 33(5):2223–2273.

Gu, S., Kelly, B. T., and Xiu, D. (2019). Autoencoder asset pricing models.

Haddad, V., Kozak, S., and Santosh, S. (2020). Factor timing. The Review of Financial Studies, 33(5):1980–2018.

Harvey, C. R. and Siddique, A. (2000). Conditional skewness in asset pricing tests. The Journal of finance, 55(3):1263–1295.

Henkel, S. J., Martin, J. S., and Nardari, F. (2011). Time-varying short-horizon predictability. Journal of financial economics, 99(3):560–580.

Heston, S. L. and Sadka, R. (2008). Seasonality in the cross-section of stock returns. Journal of Financial Economics, 87(2):418–445.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8):1735–1780.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. The Review of Financial Studies, 28(3):650–705.

Jensen, T. I., Kelly, B. T., and Pedersen, L. H. (2022). Is there a replication crisis in finance? Journal of Finance, forthcoming.

Kingma, D. P. and Ba, J. (2015). Adam: A method for stochastic optimization. In Bengio, Y. and LeCun, Y., editors, 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings.

Lynch, A. W. (2001). Portfolio choice and equity characteristics: Characterizing the hedging demands induced by return predictability. Journal of Financial Economics, 62(1):67–130.

Markowitz, H. (1952). The utility of wealth. Journal of political Economy, 60(2):151–158.

Moreira, A. and Muir, T. (2017). Volatility-managed portfolios. The Journal of Finance, 72(4):1611–1644.

Moreira, A. and Muir, T. (2019). Should long-term investors time volatility? Journal of Financial Economics, 131(3):507–527.

Polk, C., Vayanos, D., and Woolley, P. (2022). Long-horizon investing in a non-capm world. Available at SSRN 4096829.

Simonyan, K., Vedaldi, A., and Zisserman, A. (2013). Deep inside convolutional networks: Visualising image classification models and saliency maps. arXiv preprint arXiv:1312.6034.

Stambaugh, R. F. (1999). Predictive regressions. Journal of financial economics, 54(3):375–421.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems, pages 5998–6008.

Welch, I. and Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. The Review of Financial Studies, 21(4):1455–1508.

Zhang, C. and Aaraba, A. (2022). Dynamic optimal portfolio with trading cost constraints via deep reinforcement learning. Working Paper, McGill University.

A Appendix

A.1 Stocks-Specific Features

Table 1: Stock-specific Features

Feature	Acronym	Reference
Firm age	age	Jiang Lee and Zhang (2005)
Liquidity of book assets	aliq_at	Ortiz-Molina and Phillips (2014)
Liquidity of market assets	aliq_mat	Ortiz-Molina and Phillips (2014)
Amihud Measure	ami_126d	Amihud (2002)
Book leverage	at_be	Fama and French (1992)
Asset Growth	at_gr1	Cooper Gulen and Schill (2008)
Assets-to-market	at_me	Fama and French (1992)
Capital turnover	at_turnover	Haugen and Baker (1996)
Change in common equity	be_gr1a	Richardson et al. (2005)
Book-to-market equity	be_me	Rosenberg Reid and Lanstein (1985)
Market Beta	beta_60m	Fama and MacBeth (1973)
Dimson beta	beta_dimson_21d	Dimson (1979)
Frazzini-Pedersen market beta	betabab_1260d	Frazzini and Pedersen (2014)
Downside beta	betadown_252d	Ang Chen and Xing (2006)
Book-to-market enterprise value	bev_mev	Penman Richardson and Tuna (2007)
The high-low bid-ask spread	bidaskhl_21d	Corwin and Schultz (2012)
Abnormal corporate investment	capex_abn	Titman Wei and Xie (2004)
CAPEX growth (1 year)	capx_gr1	Xie (2001)
CAPEX growth (2 years)	capx_gr2	Anderson and Garcia-Feijoo (2006)
CAPEX growth (3 years)	capx_gr3	Anderson and Garcia-Feijoo (2006)
Cash-to-assets	cash_at	Palazzo (2012)
Net stock issues	chcsho_12m	Pontiff and Woodgate (2008)
Change in current operating assets	coa_gr1a	Richardson et al. (2005)
Change in current operating liabilities	col_gr1a	Richardson et al. (2005)
Cash-based operating profits-to-book assets	cop_at	
Cash-based operating profits-to-lagged book assets	cop_atl1	Ball et al. (2016)
Market correlation	corr_1260d	Assness, Frazzini, Gormsen, Pedersen (2020)
Coskewness	coskew_21d	Harvey and Siddique (2000)
Change in current operating working capital	cowc_gr1a	Richardson et al. (2005)
Net debt issuance	dbnetis_at	Bradshaw Richardson and Sloan (2006)
Growth in book debt (3 years)	debt_gr3	Lyandres Sun and Zhang (2008)
Debt-to-market	debt_me	Bhandari (1988)
Change gross margin minus change sales	dgp_dsale	Abarbanell and Bushee (1998)
Dividend yield	div12m_me	Litzenberger and Ramaswamy (1979)
Dollar trading volume	dolvol_126d	Brennan Chordia and Subrahmanyam (1998)
Coefficient of variation for dollar trading volume	dolvol_var_126d	Chordia Subrahmanyam and Anshuman (2001)
Change sales minus change Inventory	dsale_dinv	Abarbanell and Bushee (1998)
Change sales minus change receivables	dsale_drec	Abarbanell and Bushee (1998)
Change sales minus change SG&A	dsale_dsga	Abarbanell and Bushee (1998)
Earnings variability	earnings_variability	Francis et al. (2004)
Return on net operating assets	ebit_bev	Soliman (2008)
Profit margin	ebit_sale	Soliman (2008)
Ebitda-to-market enterprise value	ebitda_mev	Loughran and Wellman (2011)
Hiring rate	emp_gr1	Belo Lin and Bazdresch (2014)

Table 1 continued from previous page

Feature	Acronym	Reference
Equity duration	eq_dur	Dechow Sloan and Soliman (2004)
Net equity issuance	eqnetis_at	Bradshaw Richardson and Sloan (2006)
Equity net payout	eqnpo_12m	Daniel and Titman (2006)
Net payout yield	eqnpo_me	Boudoukh et al. (2007)
Payout yield	eqpo_me	Boudoukh et al. (2007)
Pitroski F-score	f_score	Piotroski (2000)
Free cash flow-to-price	fcf_me	Lakonishok Shleifer and Vishny (1994)
Change in financial liabilities	fnl_gr1a	Richardson et al. (2005)
Gross profits-to-assets	gp_at	Novy-Marx (2013)
Gross profits-to-lagged assets	gp_atl1	
Intrinsic value-to-market	intrinsic_value	Frankel and Lee (1998)
Inventory growth	inv_gr1	Belo and Lin (2011)
Inventory change	inv_gr1a	Thomas and Zhang (2002)
Idiosyncratic skewness from the CAPM	iskew_capm_21d	
Idiosyncratic skewness from the Fama-French 3-factor model	iskew_ff3_21d	Bali Engle and Murray (2016)
Idiosyncratic skewness from the q-factor model	iskew_hxz4_21d	
Idiosyncratic volatility from the CAPM (21 days)	ivol_capm_21d	
Idiosyncratic volatility from the CAPM (252 days)	ivol_capm_252d	Ali Hwang and Trombley (2003)
Idiosyncratic volatility from the Fama-French 3-factor model	ivol_ff3_21d	Ang et al. (2006)
Idiosyncratic volatility from the q-factor model	ivol_hxz4_21d	
Kaplan-Zingales index	kz_index	Lamont Polk and Saa-Requejo (2001)
Change in long-term net operating assets	lnoa_gr1a	Fairfield Whisenant and Yohn (2003)
Change in long-term investments	lti_gr1a	Richardson et al. (2005)
Market Equity	market_equity	Banz (1981)
Mispricing factor: Management	mispricing_mgmt	Stambaugh and Yuan (2016)
Mispricing factor: Performance	mispricing_perf	Stambaugh and Yuan (2016)
Change in noncurrent operating assets	ncoa_gr1a	Richardson et al. (2005)
Change in noncurrent operating liabilities	ncol_gr1a	Richardson et al. (2005)
Net debt-to-price	netdebt_me	Penman Richardson and Tuna (2007)
Net total issuance	netis_at	Bradshaw Richardson and Sloan (2006)
Change in net financial assets	nfna_gr1a	Richardson et al. (2005)
Earnings persistence	ni_ar1	Francis et al. (2004)
Return on equity	ni_be	Haugen and Baker (1996)
Number of consecutive quarters with earnings increases	ni_inc8q	Barth Elliott and Finn (1999)
Earnings volatility	ni_ivol	Francis et al. (2004)
Earnings-to-price	ni_me	Basu (1983)
Quarterly return on assets	niq_at	Balakrishnan Bartov and Faurel (2010)
Change in quarterly return on assets	niq_at_chg1	
Quarterly return on equity	niq_be	Hou Xue and Zhang (2015)
Change in quarterly return on equity	niq_be_chg1	
Standardized earnings surprise	niq_su	Foster Olsen and Shevlin (1984)
Change in net noncurrent operating assets	nncoa_gr1a	Richardson et al. (2005)
Net operating assets	noa_at	Hirshleifer et al. (2004)
Change in net operating assets	noa_gr1a	Hirshleifer et al. (2004)
Ohlson O-score	o_score	Dichev (1998)
Operating accruals	oaccruals_at	Sloan (1996)
Percent operating accruals	oaccruals_ni	Hafzalla Lundholm and Van Winkle (2011)
Operating cash flow to assets	ocf_at	Bouchard, Krüger, Landier and Thesmar (2019)
Change in operating cash flow to assets	ocf_at_chg1	Bouchard, Krüger, Landier and Thesmar (2019)

Table 1 continued from previous page

Feature	Acronym	Reference
Operating cash flow-to-market	ocf_me	Desai Rajgopal and Venkatachalam (2004)
Cash flow volatility	ocfq_saleq_std	Huang (2009)
Operating profits-to-book assets	op_at	
Operating profits-to-lagged book assets	op_atl1	Ball et al. (2016)
Operating profits-to-book equity	ope_be	Fama and French (2015)
Operating profits-to-lagged book equity	ope_bel1	
Operating leverage	opex_at	Novy-Marx (2011)
Taxable income-to-book income	pi_nix	Lev and Nissim (2004)
Change PPE and Inventory	ppeinv_gr1a	Lyandres Sun and Zhang (2008)
Price per share	prc	Miller and Scholes (1982)
Current price to high price over last year	prc_highprc_252d	George and Hwang (2004)
Quality minus Junk: Composite	qmj	Assness, Frazzini and Pedersen (2018)
Quality minus Junk: Growth	qmj_growth	Assness, Frazzini and Pedersen (2018)
Quality minus Junk: Profitability	qmj_prof	Assness, Frazzini and Pedersen (2018)
Quality minus Junk: Safety	qmj_safety	Assness, Frazzini and Pedersen (2018)
R&D-to-market	rd_me	Chan Lakonishok and Sougiannis (2001)
R&D-to-sales	rd_sale	Chan Lakonishok and Sougiannis (2001)
R&D capital-to-book assets	rd5_at	Li (2011)
Residual momentum t-12 to t-1	resff3_12_1	Blitz Huij and Martens (2011)
Residual momentum t-6 to t-1	resff3_6_1	Blitz Huij and Martens (2011)
Short-term reversal	ret_1_0	Jegadeesh (1990)
Price momentum t-12 to t-1	ret_12_1	Fama and French (1996)
Price momentum t-12 to t-7	ret_12_7	Novy-Marx (2012)
Price momentum t-3 to t-1	ret_3_1	Jegadeesh and Titman (1993)
Price momentum t-6 to t-1	ret_6_1	Jegadeesh and Titman (1993)
Long-term reversal	ret_60_12	De Bondt and Thaler (1985)
Price momentum t-9 to t-1	ret_9_1	Jegedeesh and Titman (1993)
Maximum daily return	rmax1_21d	Bali Cakici and Whitelaw (2011)
Highest 5 days of return	rmax5_21d	Bali, Brown, Murray and Tang (2017)
Highest 5 days of return scaled by volatility	rmax5_rvol_21d	Assness, Frazzini, Gormsen, Pedersen (2020)
Total skewness	rskew_21d	Bali Engle and Murray (2016)
Return volatility	rvol_21d	Ang et al. (2006)
Assets turnover	sale_bev	Soliman (2008)
Labor force efficiency	sale_emp_gr1	Abarbanell and Bushee (1998)
Sales Growth (1 year)	sale_gr1	Lakonishok Shleifer and Vishny (1994)
Sales Growth (3 years)	sale_gr3	Lakonishok Shleifer and Vishny (1994)
Sales-to-market	sale_me	Barbee Mukherji and Raines (1996)
Sales growth (1 quarter)	saleq_gr1	
Standardized Revenue surprise	saleq_su	Jegadeesh and Livnat (2006)
Year 1-lagged return, annual	seas_1_1an	Heston and Sadka (2008)
Year 1-lagged return, nonannual	seas_1_1na	Heston and Sadka (2008)
Years 11-15 lagged returns, annual	seas_11_15an	Heston and Sadka (2008)
Years 11-15 lagged returns, nonannual	seas_11_15na	Heston and Sadka (2008)
Years 16-20 lagged returns, annual	seas_16_20an	Heston and Sadka (2008)
Years 16-20 lagged returns, nonannual	seas_16_20na	Heston and Sadka (2008)
Years 2-5 lagged returns, annual	seas_2_5an	Heston and Sadka (2008)
Years 2-5 lagged returns, nonannual	seas_2_5na	Heston and Sadka (2008)
Years 6-10 lagged returns, annual	seas_6_10an	Heston and Sadka (2008)
Years 6-10 lagged returns, nonannual	seas_6_10na	Heston and Sadka (2008)
Change in short-term investments	sti_gr1a	Richardson et al. (2005)
Total accruals	taccruals_at	Richardson et al. (2005)
Percent total accruals	taccruals_ni	Hafzalla Lundholm and Van Winkle (2011)

Table 1 continued from previous page

Feature	Acronym	Reference
Asset tangibility	tangibility	Hahn and Lee (2009)
Tax expense surprise	tax_gr1a	Thomas and Zhang (2011)
Share turnover	turnover_126d	Datar Naik and Radcliffe (1998)
Coefficient of variation for share turnover	turnover_var_126d	Chordia Subrahmanyam and Anshuman (2001)
Altman Z-score	z_score	Dichev (1998)
Number of zero trades with turnover as tiebreaker (6 months)	zero_trades_126d	Liu (2006)
Number of zero trades with turnover as tiebreaker (1 month)	zero_trades_21d	Liu (2006)
Number of zero trades with turnover as tiebreaker (12 months)	zero_trades_252d	Liu (2006)

B Factor Timing

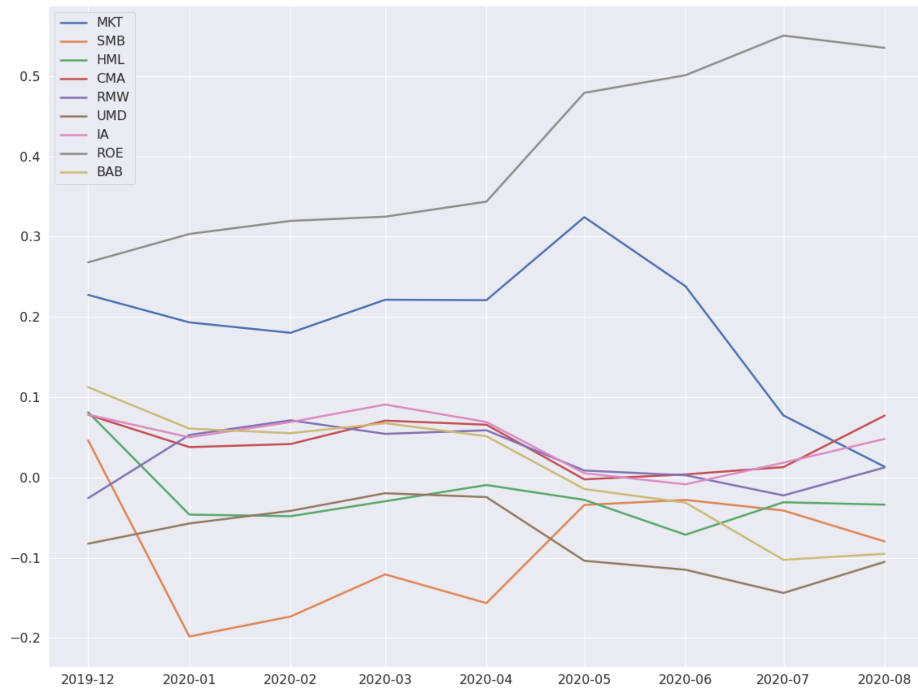
B.1 COVID-19 Market Crush

As a special case we isolate the year 2020, and focus on how monthly rebalancing, and 12 month investment horizon RL strategy allocates the weights around COVID-19 March, 2020 market plunge. 2020 is the last year of our *OOS* period, and the model was retrained last time by the end of 2019. Moreover, the model has never been trained on anything similar to this pandemic episode as it never happens before in our sample period. This event therefore provides a unique laboratory experiment to examine how after being trained on previous crisis episodes, the model makes decision for something it has never experienced before.

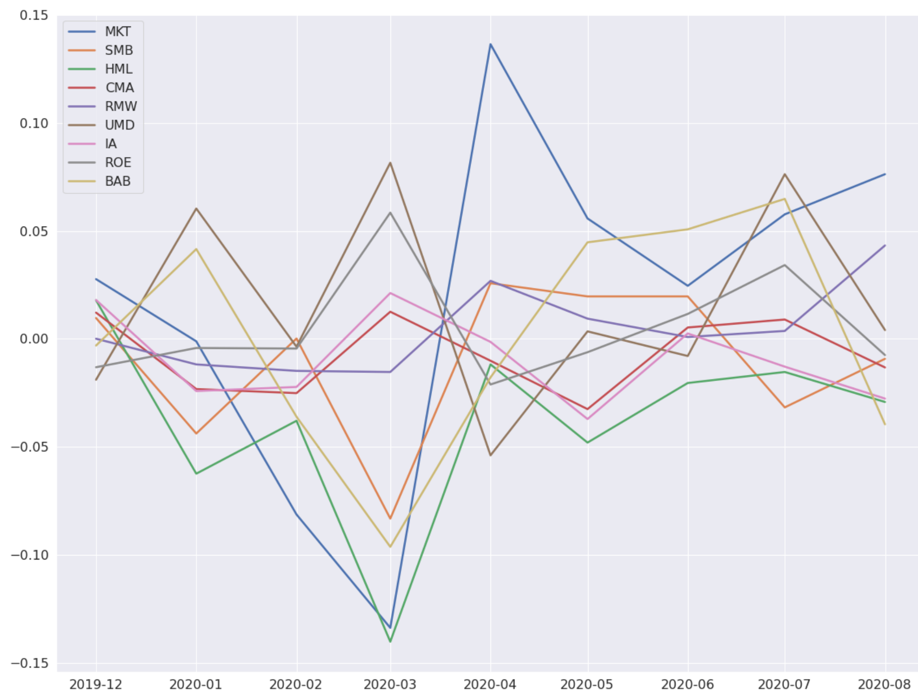
Figure 1a plots monthly weights for all 9 factors during the year 2020, while Figure 1b plots monthly factor returns to be able to relate their performances to the ex-ante weights allocations.

MKT starts declining in January 2020, with the biggest drop in March, Figure 1b. The model never short-sells MKT, but rather decreases its weight from January to February, and then increase it again to January level in March, at the market bottom, Figure 1a. It keeps the weight of March through April, when the market picks up the most, i.e. it does not purchase more at the pick. It further increases the weight of MKT in May, when the market retreats from April's high. These *ex-ante* model decisions can be found quite rational and effective *ex-post* by institutional, i.e. pension funds, portfolio managers who cannot short-sell the market, but can time the market volatility (Moreira and Muir (2019)).

Among other factors, HML and SMB drop significantly in March and both finish the year with negative returns. The model keeps negative position in both HML and SMB through the year. While HML weights are relatively stable during this year, the model starts short-selling SMB in January 2020, and from about 5% December 2019 weight it decreases it to -20%.



(a) Dynamic Weights Allocation by RL Model



(b) Monthly Factors' Returns Realizations

Figure 1: Year 2020, COVID-19 Market Plunge and Recovery: dynamic weights allocation by RL model and monthly factors' returns realizations

Therefore, instead of shorting the market, MKT, it shorts the positive exposure to the small market cap, which during market downturns is affected more. Note, that the weights plotted in Figure 1a are ex-ante, i.e. the weight for January is to enter into this position either in the end of December 2019, or at the very beginning of January 2020. Later in March the model partially covers its short position with the new weight about -12%. This is also the time when the short-position is the most in-the-money. As the market begins to rebound in April, the model decreases this short position even further. These ex-ante decisions can again be found quite optimal ex-post.

Another interesting observations are factor realizations of ROE and UMD in March 2020, which went opposite to MKT, and spiked instead. The model continues shorting UMD as it never really recovered after 2008-2009 financial crisis. It already kept high weight in ROE in January, 30%, and it only starts increasing ROE weight in April, after it dropped from March high. The preferential choice of ROE by the trained model is not surprising, given the volatility penalty in our objective function, eq. 9. ROE does not drop as much as other factors during market downturns, Figure 7b, and continues exhibiting upward trend through the whole sample period.

Overall, the model seems to detect the ex-ante risks correctly, even in the environment it has never been trained for, and the performance is driven by positive realizations of risk premiums. Moreover, the model also relies heavily on dynamic, time-varying diversification among the factors.

B.2 2008-2009 Financial Crisis

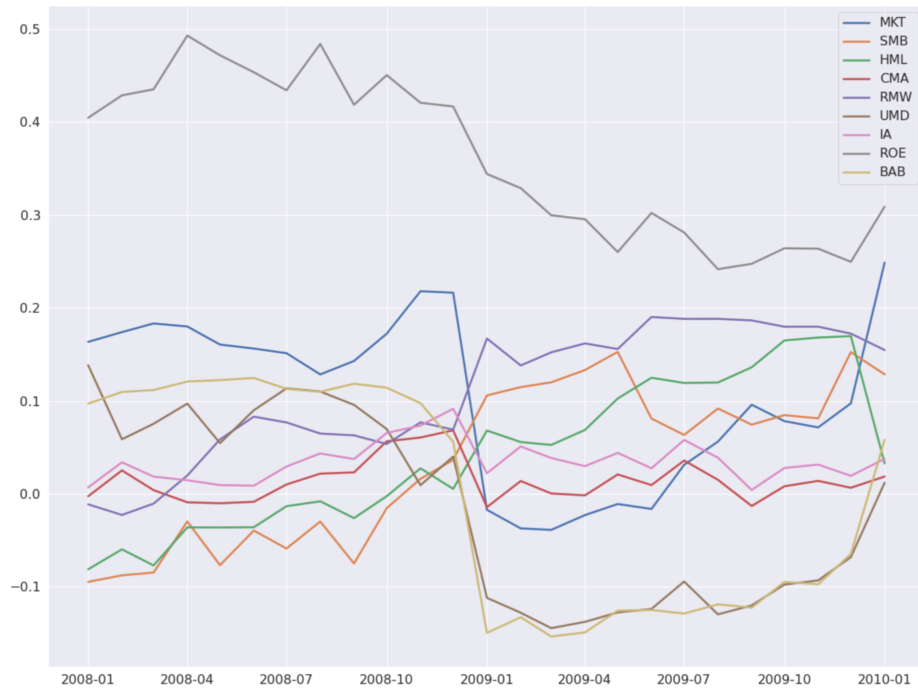
As the final isolated event analysis, we examine how the model allocates weights during the great financial 2008-2009 crisis. This episode too falls in our post-2005 out-of-sample period, where the model has never been trained on the consequences of "too big to fail" failure. Lehman Brothers ceased all operations by the mid of September 2008, sending the market into the turmoil and panic. The market plunge continued through October 2008, reaching the bottom.

Figure 2b presents monthly factor return realizations for 2008-2009. MKT from rebounding in August 2008 exhibits continues decline through the end of October 2008. The model's weighs allocations for MKT, Figure 2a do not follow August rebound, and rather apposite, it assigns to MKT the smallest for the calendar year, 2008, weight in the beginning of August - it recommends to decrease holdings in the market. For September, the first month of decline, the model recommends buying MKT, and further increasing the position in MKT for October, the ex post realized market bottom. It stops increasing the position in MKT for November and December of 2008. Then for January 2009, for the first time, it predicts to take a short position in MKT, which it further increases for February of 2009. Figure 2b demonstrates that the model perfectly timed the second biggest MKT plunge of January and February 2009. As the market rebounds later through the year, the model suggest the second largest positive holdings in MKT factor.

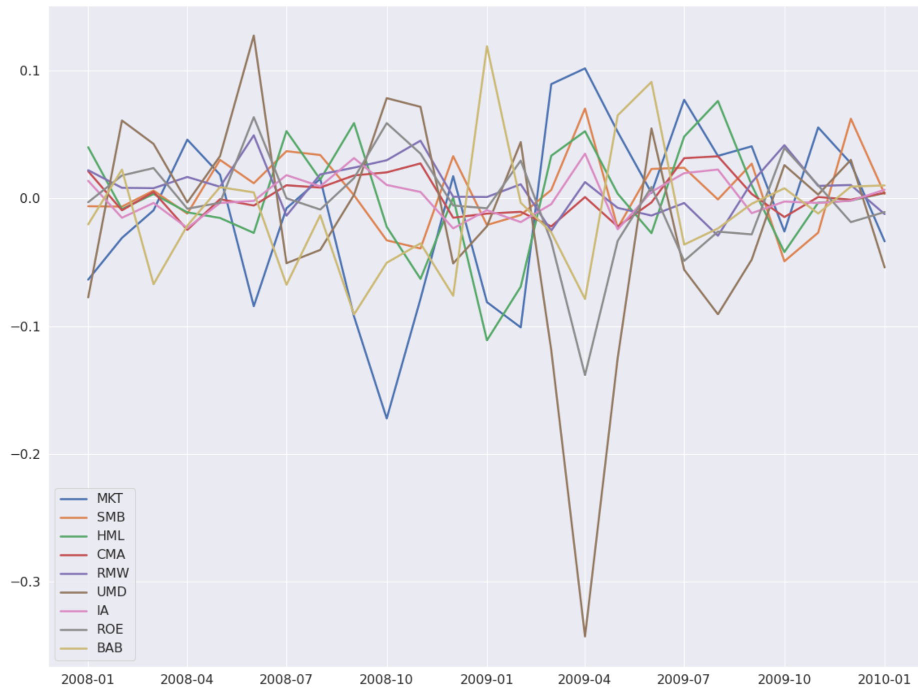
Two important observations. First, the model did not take the full advantage of September 2008 market decline, but it rather decreased its MKT exposure. The cause of September 2008 panic was a sudden bankruptcy announcement of Lehman Brothers, which obviously was a systemic shock rather than the market shock. This shock was not only outside of any financial data grasp, it was also outside of "too big to fail" perception grasp which dominated the market sentiment at that time. Yet, the model predicted in the right direction. Second, it perfectly timed the second market plunge of January and February 2009, since it was a market shock, which can be captured and predicted by financial data provided to the model.

UMD has the highest decline among all factors in April 2009, followed by substantially smaller declines of ROE and BAB at the same time, Figure 2b. Figure 2a shows that the model recommends a short position in UMD and BAB far before April, for January 2009. Here it again exhibits a perfect timing of factor predictability and aggressive short positions. This dynamics also illustrates how the RL portfolio manages to "escape" the deepest MKT cumulative decline in early 2009, Figure 7a.

Overall, the factor predictability and factor premium realizations timing of our RL model is consistent with the evidence of Haddad et al. (2020). The authors too show that factors can be predicted, and the timing strategies of factors based on this predictability can provide substantial economic gains. We support their results by demonstrating even larger economic gains from managing multi-factor portfolios, i.e. extracting further extra-gains from diversification effect among factors on a portfolio level in one consistent dynamic framework.



(a) Dynamic Weights Allocation by RL Model



(b) Monthly Factors' Returns Realizations

Figure 2: 2008-2009 Financial Crisis: dynamic weights allocation by RL model and monthly factors' returns realizations