Keeping the Faith (and the Returns): An AI Approach to Values-based Investing*

Maureen O'Hara
Johnson College of Business, Cornell University
mo19@cornell.edu

Artem Streltsov
SUNY Buffalo
artemstr@buffalo.edu

June 2025

Abstract

We demonstrate a new approach for SRI investing drawing on Large Language Models (LLMs) and Deep Reinforcement Learning (DRL) to address the challenges posed by moral constraints on portfolio selection. Using the Global X S&P500 CATH ETF (the largest Catholic values fund) as an example, we use textual analysis to identify companies consistent with the values mandate, allowing us to create "synthetic CATH" portfolios of different sizes and with longer time horizons. We further optimize each portfolio using DRL to arrive at an optimal set of portfolio weights that maximize out-of-sample Sharpe ratios. We demonstrate dramatic improvements in risk-adjusted returns closing the performance gap in values-based investing.

JEL classification: G11, G12, G24

Keywords: SRI, Faith-based investing, artificial intelligence, large language models, machine learning, reinforcement learning

* We thank Reena Aggarwal, Will Cong, Fr. Quentin Dupont, David Easley, Will Goetzman, John Griffin, Xinwei Liu, Sydney Ludvigson, Jose Minaya, Raghavendra Rao, Laura Starks and seminar participants at Cornell and Hong Kong University for helpful comments.

Keeping the Faith (and the Returns): An AI Approach to Values-based Investing

1. Introduction

Socially responsible investing (SRI, also known as ESG) falls into two categories – value-based and values-based (Starks (2023)). Value-based investing incorporates into portfolio selection the risk of holding particular types of assets, adjusting the holding size and return expectation (e.g. green investing)[1]. Values-based investing incorporates moral views into asset selection, with restrictions on what assets are eligible for inclusion into a portfolio (e.g. faith-based investing). While much of the attention in SRI has focused on the first category, values-based investing has been growing, both due to social and politically-based divestment mandates and to increasing interest in faith-based investing.[2] Screening out ineligible stocks that misalign with an investor's values, however, places additional constraints on the portfolio optimization problem. This can lead, all else equal, to a suboptimal trading strategy that trades off portfolio returns for values-conscious investment activity. Moreover, even determining these ineligible stocks poses difficulties for fund managers who have to translate moral-based preferences into specific stock selections. In this paper we propose a framework to rekindle the faith in values-based investment using state-of-the-art machine learning techniques.

To motivate why such an artificial intelligence approach can be fruitful, consider the following. Values-based investing often requires discerning eligible investments where the metrics defining the constraints may not be unequivocal. Identifying these firms with simple numerical metrics may fail to capture the needed moral dimensions of the investment. Moreover, many of the largest U.S. values-based funds and indices share several common characteristics.[3] They are relatively young and thus have limited trading history. Such funds also have limited

[1] This focus on risk is captured in Blackrock's 2020 letter to clients stating that the firm would henceforth consider "ESG risk with the same rigor that it analyzes traditional measures such as a credit risk and liquidity risk" (see discussion in Matos (2020) and "Sustainability as Blackrock's New Standard for Investing", 14 January 2020).
[2] Overall size estimates of faith-based investing are difficult to determine. A recent report by Morningstar (2023) counted 850 faith-based funds and ETFs globally, with combined assets of $100 B, split between Christian and Shariah funds. Most of the assets are Christian-based, with 95% of such funds based in the US. The Interfaith Center for Corporate Responsibility includes 300 global institutional investors with more than $4 trillion in managed assets.

[3] The largest global provider of faith-based ETFs is Inspire Investing with $1.4 billion in ETFs such as BIBL, BLES, and GLRY. The largest faith-based fund is the $5.3 billion Amana Growth Investor (AMAGX) which focuses on Shariah-compliant investments. This fund holds a mix of equities and bonds.

scope and focus on only, say, S&P 500 firms. Most importantly, they value-weight their constituents rather than optimize the holding weights using some model. These latter features speak to restrictions having little to do with moral preferences but much to do with limiting potential returns. In this paper we ask: can we do better?

To this end, we use state-of-the-art Natural Language Processing (NLP) techniques, most notably Large Language Models (LLMs), to uncover what kind of firms align with the values in question. Using the tools of textual analysis, we can learn both from values-based fund managers' revealed preferences in stock selection and from discerning moral dimensions of firms from earnings conference calls. This lets us construct a synthetic fund consistent with these values, but with the longer trading history needed to build better algorithms and to validate trading strategies with higher precision. We can also expand beyond S&P 500 firms to construct values-based portfolios of different sizes taking advantage of opportunities in smaller firms that did not happen to be in the S&P 500.

We further optimize each portfolio using Deep Reinforcement Learning (DRL) to arrive at an optimal set of portfolio weights that maximize out-of-sample Sharpe ratios. Unlike supervised and unsupervised machine learning approaches, reinforcement learning uses a trial-and-error approach and a policy to learn dynamically.[4] This allows the algorithm to choose a particular portfolio design, learn from the market reaction, recommend the next optimal design, and so on. Deep reinforcement learning uses neural networks to approximate the policy function, thereby allowing higher-dimensional state spaces, facilitating the ability to learn more complex strategies and adapt to a changing environment in real time.

Our specific focus in this paper is on faith-based investing, but our broader contribution is to develop a methodology that can be applied to a wide range of settings. In particular, we believe our AI-based methodology is applicable to portfolio optimization problems subject to textual, nebulous, or implicit constraints. Broader SRI mandates such as climate or social impact can fall into this category, as can restrictions on portfolios arising from tax mandates, or geographic limitations, or simply thematic preferences such as sentiment. In these, and many other

[4] The specific class of RL algorithms we base our discussion on and implement in this paper is Policy Gradient Algorithms that suit our setting best – learning an optimal action-state mapping with a continuous action space via gradient ascent. Other powerful methods focus on the value function instead.

applications, restrictions on allowable assets impose constraints on portfolio selection; our AI approach provides a way to reduce this impact on portfolio return.

While faith-based investing comes in many variants, we apply our methodology to the largest Catholic-values based fund, the Global X S&P 500 Catholic Values ETF (CATH). This fund was created in April 2016, it is intended to conform to the investing guidelines given by the U.S. Council of Catholic Bishops, and it is limited to S&P 500 stocks. We demonstrate how any such ETF can be expanded to form portfolios of eligible stocks of a pre-specified size and reoptimized for social gain. Our NLP algorithm uses conference calls to classify stocks as being CATH-eligible, allowing us to create "CATHlike" synthetic portfolios with 500, 750, 1000 and 1500 stocks. In forming our synthetic portfolios, we retain the same restrictions imposed by CATH on maximum allowed individual stock weights.

With CATH holdings being limited to S&P500 universe, VTI (Vanguard Total Stock Market Index Fund ETF) and especially SPY (SPDR S&P 500 ETF Trust) are the best benchmark ETFs. For out-of-sample performance comparison, we focus on the period 2017-2022 since (1) conference call coverage is substantially better starting May of 2017; (2) CATH ETF started trading at the end of April, 2016, and (3) we have ETF holdings data from ETF Global starting January 2017. We can thus only make meaningful comparisons between value-weighted CATHlikes, SPY, VTI, CATH, optimized CATH and CATHlikes for the period limited to 2017-2022.

First, we confirm that, indeed, CATH underperforms SPY having performed measurably worse than SPY in terms of all performance metrics, while also having more negative VAR and Expected Shortfall at 5%. We further show that the degree of CATHlikeness of a stock negatively predicts future returns and a long-short portfolio going long unCATHlike stocks and short CATHlike earns a statistically and economically significant alpha demonstrating the trade-off between portfolio performance and values-based investing. The value-weighted synthetic CATHlikes constructed by our model generally performed comparably to CATH itself. Interestingly enough, large stocks have so far been doing particularly well in the 21st century making performance of value-weighted CATHlikes decreasing in portfolio size.

If, instead of value weighting, we optimize holding weights using Reinforcement learning, the optimized CATHlikes boast impressive performance with Sharpe ratios at least

44% greater than SPY or CATH. Moreover, the performance improves the more stocks are included into a portfolio, with the RL optimized CATHlike portfolio with 1500 stocks having a 100% higher Sharpe ratio than SPY. Cumulative return since 2017 of a RL optimized CATHlike 500, a portfolio most comparable to SPY, is 188 percentage points higher than that of SPY. The largest CATHlike portfolio with 1500 stocks, on the other hand, has a cumulative return 1200 percentage points higher than SPY when RL optimized.

While these expanded portfolio strategy returns are remarkably high, we also find that simply following CATH stock selection, that is keeping CATH stock composition fixed, and reweighting CATH holdings using Reinforcement Learning yields significant improvement: cumulative return is higher than both SPY and CATH by more than 75% over 2017-2022. These enhanced returns show how artificial intelligence tools can improve portfolio weightings even within restrictive selection criteria.

Finally, we address the issue of what is driving these enhanced returns. We show that these portfolios have a large Fama-French 4 and 5 alpha ranging from 11% to 70% per year depending on portfolio type and size, implying that the performance is greater than that implied by exposure to common risk factors. Because AI-based strategies often involve extensive trading, a natural concern is that ignored transaction costs are behind any outperformance. We confirm, however, that our performance gains survive even very high transaction costs. Using regression analysis, we investigate the stock characteristics underpinning our AI-selected portfolio weights, demonstrating anew why value-weighting is a such poor selection criteria. For faith-based investors, our results suggest that returns and values need not be inexorably opposed.

Our focus on values-based investing relates to what is now a voluminous body of research on the impact of environmental, social and governance issues on asset prices and returns (see, for example, research by Starks (2023); Barber, Morse and Yasuda (2021); Bolton and Kacperczyk (2021); Murfin and Spiegel (2020); Matvos (2020); Krueger, Saunter and Starks (2020)). An excellent survey of this area is Edmonds and Kacpercyzk (2022). Much of this research has focused on the risks posed by various SRI factors. A related SRI literature looks at the effects of impact investing on the cost of capital for the underlying "brown" versus "green" firms (see Pastor, Stambaugh, and Taylor (2021; 2022); Greene and Roth (2025); Berk and van Binsbergen (2021)). The problem we focus on is can you restructure a portfolio so as to offset

the missing returns from the ineligible stocks. We view the main contribution of our research as demonstrating a new approach for SRI investing that uses techniques from artificial intelligence to enhance investor returns.

Our research is also closely related to the literature on "sin stocks" which directly looks at investment restrictions motivated by moral concerns. Important papers here include Fabozzi, Ma, and Oliphant (2008); Hong and Kacperczyk (2009); Fabozzi et al (2019); and Pedersen, Fitzgibbons and Pomorski (2020). Much of this literature focusses on the trade-offs between returns and "moral" values, finding, for example, that portfolios long such precluded sin stocks outperform, while those excluding such stocks underperform.[5] Our analysis also shows such underperformance, but our demonstration that this underperformance can be mitigated (or even surmounted) at least for the Catholic values-based investment ETF shown here has important implications for the "doing good" and "doing well" tradeoff.

From a technical perspective, our research relates to a growing literature applying artificial intelligence to asset pricing. A variety of authors (notable papers here include Gu, Kelly, and Xiu (2020); Krauss, Do, and Huck (2022); Kelly, Malamud, and Zhou (2023)) have applied various machine learning techniques to portfolio selection, detecting arbitrage opportunities, and cross-sectional returns. A paper close in spirit to our work here is De Nard, Engle and Kelly (2024) who use a new approach to account for risk and "greenness" to determine factor-mimicking portfolios for hedging climate risk. An excellent survey of this literature is Kelly and Xiu (2023).

There is also a growing literature using NLP to address issues in market behavior (see, for example, Chava, Du, and Malakar (2022)), with more recent work (see Gabaix et al (2023)) using advances in NLP such as transformer-based models to extract information from more complex language models and to process ever larger data sets in practical time frames. Our NLP approach uses similar transformer-based models to extract information from industry dummies, term-frequency – inverse document frequency (TF-IDF) text representations, and features from a Large Language Model (BigBird) that learns from the market reaction to each conference call. The ultimate goal of our NLP pipeline is to detect ineligible stocks that do not align with values-

[5] An interesting paper relating machine learning to sustainable investment objectives is Yang and Yasudo (2023) who find a negative correlation between financial and impact goals.

based investing tenets and should be screened out. Our analysis shows how both NLP and Reinforcement Learning can play important roles in affecting the efficient frontier for values-based investors.

Particularly relevant for our research as well is the new research on dynamic asset pricing systems (see Koijen and Yogo (2019)) and direct alpha portfolio construction (Cong, Tang, Wang, and Zhang (2025)). This research area uses a variety of AI techniques including deep reinforcement learning, large language models and neural networks to extract underlying drivers of asset returns. Our analysis, which draws heavily on the Cong et al (2025) approach, shows how these AI tools can be used to enhance restriction-based values investing.

This paper is organized as follows. Section 2 sets out our research design, sample, and data. We also discuss the challenge of translating moral values into portfolio selections, and we compare the performance of CATH to SPY over our sample period. Section 3 determines eligible and non-eligible stocks using NLP. In training our NLP algorithm, we split our data sample into training, validation, and test sets, and evaluate its performance using standard machine learning metrics. We further form CATHlike synthetic portfolios and examine their performance relative to various benchmarks. Section 4 optimizes the weights of these portfolios, drawing on a deep reinforcement learning approach. We investigate these optimized returns relative to standard asset pricing benchmarks and compare results from our optimization schedule with well-known ETFs. We further demonstrate in Section 5 that our performance gains survive even very large transaction costs and investigate what drives the decisions of our RL agent. We conclude with a discussion of the applicability of our methodology to a range of restriction-based investment problems. The Appendix provides more information on the data and variable construction.

2. **Research design and data**

Our research goal is to create portfolios of stocks that are consistent with Catholic values and provide larger returns than are achieved by the CATH ETF. Stock selection is at the heart of any investment product, but faith-based investing imposes the additional constraint that the company's products must be aligned with underlying religious values and principles. In this section we set out the challenges involved in translating these values into portfolio selections. As

noted in the introduction, we are focusing on the particular example of Catholic faith-based investing, so we first set out the principles underlying this segment of values-based investing. We then turn to how these are applied in the largest Catholic-values based fund, the Global X S&P 500 Catholic Values ETF (CATH). We next discuss our research design of how we use machine learning techniques to select faith-consistent stocks, create synthetic CATHlike portfolios, and optimize the weights within these portfolios to enhance return. We also discuss the data sets used in our analyses.

A. Translating values into portfolio selections

Catholic faith-based investing is informed by guidance from the Vatican and the US Conference of Catholic Bishops (USCCB)[6]. The USCCB Socially Responsible Investing Guidelines (2021) sets out three strategies that should guide investing: avoid doing harm; actively work for change; and promote the common good. With these guidelines, the USCCB encourages "not investing in" or "divesting from" companies that do not: protect human life; protect human dignity; enhance the common good; pursue economic justice; or are not working to save our global common home.[7] The USCCB does not identify specific companies whose products or services fall within these proscribed areas, but they do list some specific activities (for example, adult entertainment, controversial weapons, euthanasia, abortion, child labor) that would violate church teaching. They also suggest policies such as "not investing in firms whose primary purpose is to derive revenue from gambling or the production of tobacco", or "in companies that manufacture firearms except for hunting [or….] legitimate law enforcement uses".[8]

Implementing these values into portfolio selection is not straightforward. In essence, it involves predicting which companies fall afoul of these guidelines – and by how much. The Global X S&P 500 Catholic Values Index is designed to create a benchmark consistent with Catholic values.[9] To be part of the index, a company must be part of the S&P 500, and must not

[6] For recent Vatican guidance see The Pontifical Academy of Social Sciences, 2022, "Mensuram Bonam: Faith-based Measures for Catholic Investors", available at https://www.pass.va/en/publications/other-publications/mensuram_bonam_eng.html

[7] See page 8, "Socially Responsible Investment Guidelines for the United States Conference of Catholic Bishops", November 2021.

[8] The USCCB also encourages pro-active investment in companies whose business models promote environmental goals such as the Paris Accords

[9] Global X is part of the by Mirae Asset Management Group

exceed exposure limits on particular activities. For example, a company is excluded from the index if it has any revenue from adult entertainment production, or has more than 5% of its revenue from adult entertainment distribution. A firm is excluded if more than 50% of its revenue comes from military contracting. It is not included if it has any revenue from the proscribed activities (euthanasia, etc.) noted above. The Index is value-weighed and rebalanced quarterly. The weight on an excluded firm is given to other stocks within the same GICS (Global Industries Classification Standard) sector.

The Global X S&P 500 Catholic Values ETF (CATH) is based on the index.[10] This ETF product was introduced in April 2016 and as of March 2025 had assets of $920 million and held 439 stocks. Excluded stocks generally are in consumer staples, healthcare and industrials, and include such names as Johnson & Johnson, Pfizer, Boeing, Lockheed Martin, Caesars Entertainment, and Walmart. While some of the excluded companies seem reasonable given the USCCB guidance – Lockheed Martin makes defense materials, J&J makes contraceptives, Caesars is gambling – others do not. For example, Coca Cola is included in the ETF, but PepsiCo is not. Moreover, some firms enter, leave, and then re-enter the index. A case in point being Proctor & Gamble which starts in the index in 2017, is excluded starting in early 2019, re-enters in 2021, and is currently in the top 10 holdings of CATH.[11] No explanation is given by Global X for how specific exclusions (or conversely, inclusions) are determined, but the overall goal of CATH is to be aligned with Catholic values.

We obtained holdings data monthly for CATH over the period January 2017-December 2022 from the ETF Global database.[12] As Figure 1 shows, the number of S&P 500 stocks excluded from CATH has been growing over time, and is somewhat erratic. This variability makes our research goal of discerning eligible stocks for CATH-like portfolios ever more challenging.

As might be expected, the exclusion of so many stocks can affect returns. A natural comparison benchmark is the SPDR S&P 500 Trust ETF (SPY) which is based on the S&P 500

[10] For specifics see https://www.globalxetfs.com/funds/cath/ The CATH ETF focusses on excluding stocks, and does not include a pro-active bias for stocks that do "better" in terms of meeting environmental goals. In our analysis, we follow this same approach and do not adjust for a pro-active bias.
[11] A natural concern is that this is simply a data problem, but conversations with S&P personnel indicate that this is not the case.
[12] Data on CATH holdings prior to January 2017 are not available in ETF Global.

Index. This product, like CATH, is value based and rebalanced quarterly. The two indexes are often but not always similar. CATH's most recent one-year (3-year) returns of 34.06% (8.84%) are less than SPY returns of 36.20% (10.43%).[13] A similar picture can be observed in terms of Sharpe ratios with SPY (0.64) having a larger Sharpe ratio than CATH (0.62). It is important to stress, however, that this divergence does not always arise: over our 2017- 2022 sample period, CATH outperformed SPY in 62% of months.

B. *Research design and data*

We use the firms included in CATH as a "ground truth" for what constitutes an eligible stock for a Catholic-values based portfolio. We devise an NLP algorithm to predict the probability of any stock being ineligible drawing on industry classifications, and quarterly conference calls as inputs. We develop an LLM architecture based on the BigBird encoder to produce our probability estimates. We then form CATHlike portfolios and use an extended sample period from 2004 – 2022 in our analysis of synthetic portfolios. We use reinforcement learning to optimize the weights of these CATHlike portfolios, drawing on a variety of firm specific data. We then evaluate their performance relative to common market benchmarks. We train our reinforcement learning algorithm over the period 1975 – 2003, and then test our algorithm over the period 2004 – 2022 letting the algorithm learn from its mistakes after observing each year sequentially. We also investigate whether our artificial intelligence approach can produce superior returns in portfolios restricted to the same stocks held by the CATH ETF.

One objection to our approach is that we are essentially creating an active ETF and comparing its performance to a passive ETF. We have two responses to this concern. First, active investment vehicles in general have struggled to outperform passive investments, so an improved outcome from our efforts is far from guaranteed.[14] Moreover, given that all trading decisions are made by a machine and not a person, we would argue that our approach differs from what is usually viewed as active investing. Second, the CATH ETF is the largest Catholic faith-based ETF and the portfolio construction constraints imposed there (value-based,

[13] Data are from ETF.com as of Sept. 27, 2024. Net returns will also differ due to the CATH expense ratio of .29, substantially higher than SPYs expense ratio of .09.

[14] The lackluster performance of active ETFs is highlighted in "Investment Industry Loves Active ETFs. You Probably Shouldn't", Wall Street Journal, Oct. 7, 2024. See also B. Malkiel, "Passive Investing Wins Again", Financial Times, March 22, 2025.

substitutions in adjoining GSICs) are not faith related per se. Our goal is to show that using AI we can improve on the outcomes relative to holding CATH. We do not address the interesting issue of whether, using our approach, we could improve on outcomes relative to a similarly AI optimized unconstrained portfolio.[15]

The data used in our study come from a variety of sources. We use ETF Global data to establish ETF constituent stocks over time. CRSP data provide daily, and monthly stock returns, and S&P 500 historical composition. We use Compustat data for quarterly and annual firm fundamentals. Additional data are drawn from the WRDS Financial Ratio Firm Level database. Conference call transcripts from 2002 onward are from Refinitiv. We use the Fama-French library for the Fama-French 5 factors and the risk-free rate. More information about the data and variable definitions are in the Appendix.

Finally, we raise one important caveat about our research. As noted earlier, we view the main contribution of this research as demonstrating a new approach for SRI investing that uses techniques from artificial intelligence to enhance returns. We do not view this research as promising a "free lunch" for every type of restrictive investment strategy; precluding the holdings of classes of stock may well depress returns relative to unrestricted portfolios. What our approach can do is potentially enhance these returns, reducing the trade-offs that may arise in values-based investing.

3. Determining Eligible and Ineligible Stocks using NLP

Most major values-based funds have relatively short history, but a longer trading history is needed to build better algorithms, validate our trading strategies with higher precision, and uncover long-term patterns. These funds also tend to limit themselves to subsets of the largest firms such as those included in S&P500. In this section, we explain how we use state-of-the-art Natural Language Processing (NLP) techniques to extend beyond recent history and S&P500 stocks, allowing us to take advantage of a longer training sample and of opportunities in smaller firms.

Values-based indices are commonly constructed on a screening-out basis. Companies that do not meet some predetermined criteria such as a smaller than desired percentage of revenue

[15] For example, an AI optimized SPY that allowed for varying weights as opposed to value weighting.

attributable to unwanted activities are deemed ineligible. These criteria may change over time and require an in-depth knowledge of revenue sources and product lines of each company considered. For lack of a granular dataset of this sort for every traded stock we turn to NLP to learn from fund managers' revealed preferences (stock selection) what kind of stocks are likely to be eligible for values-based investment.

The inputs to our NLP pipeline are quarterly earnings conference calls and industry definitions for each firm. Text is highly expressive and is the most descriptive data source when it comes to screening out firms based on their activities, second only to a comprehensive granular dataset of activities and revenue sources by company which are typically proprietary products of financial analytics companies. The two commonly used textual data sources in finance research are earnings conference calls and firm 10-K filings, each having their own weaknesses and strengths.

Our conference call dataset starts in 2001 and grows in coverage of firms over time (see Figure 2), while 10-K filings date as far back as the early 1990s. Description of businesses in these yearly 10-K filings is somewhat static over time, while conference calls may provide useful and variable cues through dialogue. One may suggest that managers are likely to start adapting to machines listening to them and there, indeed, is research finding some evidence in support of this (Cao et al, 2023). Managers, however, are likely to succeed by regulating their tone in conference calls rather than by withholding information since analysts and investors demand accountability and a reasonable answer addressing pressing issues, many of which are directly related to the degree such firms align with values-based investing. In this research, we focus on earnings conference calls as our input to the NLP pipeline, but in principle the analysis could also be complemented with newsfeeds, company websites or 10-K filings.

The goal of our NLP pipeline is to model the probability of an S&P500 stock being ineligible conditional on inputs with the ground truth being CATH actual holdings over time. Most stocks, around 88%, are consistent with values-based investing. We, therefore, formulate the problem as a rare event detection and choose the task of finding ineligible stocks as this is the harder task with the same end goal.

Our NLP approach combines three types of inputs of ascending complexity: industry dummies (Fama-French 48), TF-IDF (term frequency–inverse document frequency)

representation of each conference call, features from a Large Language Model that learned from the market reaction to each conference call. Each input type improves out-of-sample performance in detecting ineligible stocks.

Early versions of NLP essentially produced a simple vector with word counts from a pre-defined vocabulary and this approach has been and still is popular in finance research.[16] TF-IDF representation enhances the bag of words representation by adjusting for the frequency of words across documents, giving a larger weight to less common words and discounting less informative common words. In essence, it represents a normalized word count of each word token in the dictionary for each document. In constructing the vocabulary, we omit stop words (grammatical or function words) and explicitly require that each word appear in at least 1000 documents in total and at most in 30% of all documents. This way we exclude words that are too common and those that are too rare, arriving at a reasonable dictionary size (3000 words) and representation dimensionality. The latter directly affects the complexity of the model, but also its ability to generalize, which is especially topical given the small sample size that we are working with.

Large Language Models have revolutionized NLP. These are commonly based on some transformer architecture selected based on application. At the heart of such transformer-based models is the self-attention mechanism that produces a contextualized representation of each word token based on words in the same text passage. This mechanism bypasses the sequential dependency inherent in recurrent neural networks, the state-of-the-art in NLP prior to transformers, allowing large-scale parallelization and consequently much greater model complexity, i.e. ever larger language models, and ability to process very large datasets in practical timeframes. These two virtues paired with the self-supervised nature of language gave rise to a plethora of powerful transformers-based pretrained models and have brought transformers to the forefront of the NLP research.

A major setback of vanilla transformers is that the memory and computational complexity of the full self-attention mechanism is quadratic in the sequence length making them computationally intractable in many applications requiring concurrent analysis of longer text passages such as conference calls. Most conference call transcripts are considerably longer than

[16] The seminal paper introducing signed LM word lists in finance has nearly 6000 citations (Loughran and McDonald, 2011).

512-word tokens, the maximum input length for popular transformer-based models such as BERT. More specifically, the average number of word tokens is 1200 and 1700 for introduction and Q&A sections respectively, while roughly 99% of sections, introduction and Q&A each, have less than 4096 tokens. We take advantage of the latest developments in NLP and opt for a next-generation transformer architecture, BigBird, that approximates the self-attention matrix reducing the quadratic dependency to linear. BigBird is able to support up to 4096 word tokens in a single sequence which is enough to fit full conference call sections without losing too much information.

We pass the management introduction and the Q&A section of each conference call separately in parallel through the same BigBird model, i.e. the BigBird encoder is shared by the two sections, and further send the representation returned by the transformer through another deep neural network predicting cumulative abnormal returns around the event. The architecture of the model is depicted in Figure 3. All layers of the model, including the transformer BigBird encoder, are trained together. The neural network constructs a set of common conference call derived asset pricing factors with each conference call having a different exposure to each such common factor. These exposures, the 16×1 representations from the Deep Neural Network (DNN) hidden layer, for each conference call are our final type of input. Such layers are referred to as hidden or latent as they do not directly interact with either the inputs, the conference call text in our case, or the outputs, the market reaction.

Mathematically speaking, given the cumulative 1-day abnormal return around conference call i denoted as y_i and 1536-dimensional representation of each conference call returned by the BigBird encoder denoted as $x_i \in \mathbb{R}^{1536\times 1}$, we find, together with optimal BigBird parameters collectively denoted as BB_{params} for simplicity, two weight matrices, $W_1 \in \mathbb{R}^{16\times 1536}$ and $W_2 \in \mathbb{R}^{1\times 16}$, and two bias vectors (similar to regression intercepts they re-center the left-hand side variables), $b_1 \in \mathbb{R}^{16\times 1}$ and $b_2 \in \mathbb{R}^{1\times 1}$, such that

$$\left(W_1, W_2, b_1, b_2, BB_{params}\right) \in argmin \frac{1}{n}\sum_i (y_i - \hat{y_i})^2 \ \ where$$

$$a_i^1 = \max\left(0, W_1 x_i + b_1\right)$$

$$\hat{y_i} = a_i^1 W_2 + b_2$$

The 16-dimensional exposures used as inputs in our NLP algorithm are the activations a_i^1 defined above.

We use an L2-regularized logistic regression to produce a probability estimate of a stock being ineligible in quarter t. Regularization strength is a hyperparameter that needs to be tuned. Additionally, we need to determine the threshold on the predicted probability above which we consider a stock ineligible: $ineligible\ if\ P(ineligible) \geq threshold$. A lower threshold would make the algorithm flag more stocks as ineligible or equivalently consider eligible only stocks unlikely to be ineligible. A lower threshold thus produces a more conservative set of eligible stocks.

We decide on these hyperparameters and train the model as follows. We first split the sample into training (2017-2018), validation (2019) and test sets (2020-2021) using S&P500 stocks. We determine the optimal regularization strength hyperparameter via 5-fold cross-validation on the training set. Further using the optimal regularization hyperparameter we evaluate on the validation set and choose a threshold that maximizes validation F1 score (a harmonic average between recall and precision, i.e. what share of truly ineligible stocks are detected by an algorithm and what share of stocks flagged by an algorithm as ineligible are indeed so). Having chosen both the threshold and the regularization strength, we perform a final evaluation on the test set and produce ineligibility probabilities for all stocks in the sample since 2003.

We track several performance metrics in evaluating the power of our model out-of-sample. Accuracy is the most basic metric measuring the percentage of stocks labelled correctly as either eligible or not. Our dataset is highly imbalanced due to most stocks being eligible for CATH investing which makes accuracy an unreliable metric: a naïve and completely useless model classifying all stocks as eligible would automatically have a high accuracy. For this reason, we focus on measures more sensitive to class imbalances such as precision, share of correctly predicted ineligible stocks among all stocks predicted as such, recall, share of correctly predicted ineligible stocks relative to the total number of truly ineligible stocks, and their harmonic average, F1-score.

The Precision-Recall curve is one canonical visualization used in machine learning classification problems. The curve is generally concave, and a perfect classifier has a recall and

precision of 1 implying that the area under the Precision-Recall curve, termed the average precision, equals 1. The curve is traced by varying the confidence level threshold for classifying a stock as ineligible. In choosing the threshold, one can either opt for the one corresponding to an optimal combination of precision and recall as defined by their F1-score or one may want to minimize an error of a specific type – false positives or false negatives. The area under the Precision-Recall curve is termed Average Precision and is the most fitting metric for our setting. An alternative graphic, a Receiver Operating Curve, compares directly true positive and false positive rates. The area under this curve, ROC-AUC, is a commonly used measure summarizing how well a model is able to distinguish positive examples from negative at different confidence levels and has a maximum value of 1. A higher value of all the above-mentioned measures (precision, recall, F1-score, average precision, ROC-AUC) indicates a better model.

Quantitatively speaking, at the threshold level maximizing validation set F1-score, 80% model confidence level, our model's test set performance substantially exceeds that of the naïve baseline (random chance): accuracy of 93% against 88% for a random chance baseline, F1-score of 0.70 against 0.21, precision of 0.82, recall of 0.62, ROC-AUC of 0.88 against 0.5 and average precision of 0.75 against 0.12. Using this set of hyperparameters, we further refit the model on the full dataset and predict ineligibility of stocks outside of S&P500 and of stocks before 2017 when the CATH holdings data start.

A common misconception is that screening out unCATHlike stocks can be easily done by simply excluding certain industries. To address this claim, in Table 1 we compare out-of-sample performance of the naïve baseline with that of a Fama-French 48 logistic regression model and that of our model. The industry dummy model nearly triples average precision over the baseline demonstrating that some industries are, indeed, much more likely to be unCATHlike than others. Our model, however, that adds in conference call data increases average precision by 525% over the baseline, more than doubling this metric over the industry dummy alone model. Industry definitions alone are thus insufficient for any credible analysis of CATH stock eligibility.

Further, we use our measure of stock CATHlikeness, the probability of a stock being eligible for inclusion into CATH produced by our model, to demonstrate that faith-based investing, indeed, underperforms. This evidence emphasizes the trade-off between investing well and investing for good. More specifically, every quarter we sort stocks into quintiles based on their

CATH ineligibility and within each quintile form a value-weighted portfolio since 2004. We also construct a long-short portfolio whereby we go long the least CATHlike quintile portfolio and short the most CATHlike quintile portfolio. We further compare excess returns and Fama-French 4 alphas across portfolios.

As follows from Table 2, each portfolio has a statistically significant and positive excess return, but this return monotonically decreases in the degree of CATHlikeness with average monthly excess return ranging from 0.89% for least CATHlike stocks to 0.64% for most CATHlike stocks. In addition, a long-short portfolio going long the least CATHlike quintile portfolio and short the most CATHlike has a statistically significant average monthly return of 0.25%. In terms of annualized Fama-French 4 alphas, the least CATHlike portfolio has a statistically significant alpha of 2.87% per year, while the long-short portfolio has a statistically significant alpha of 3.59% per year. We have thus established that least CATHlike stocks are considered riskier and command a greater premium or alternatively that CATHlike stocks underperform.

Armed with quarterly stock-level probabilities of CATHlikeness, we next construct an array of CATHlike portfolios since 2004 with 500, 750, 1000 and 1500 stocks. Conference call coverage increases over time such that only after 2007, 2011 and 2017 are there enough stocks to form proper 750-stock, 1000-stock and 1500-stock eligible portfolios. Interestingly enough, this conference call coverage issue is not limited to small stocks with many S&P500 stocks also affected especially in early years. Conference call coverage gets wider from 2017 onward.

Fortunately, since CATH inception we have enough data to form proper portfolios of all sizes. We require for each stock-month at least 80% of valid values for our feature set used for Reinforcement Learning optimization purposes (the list of predictors is covered in the appendix) and impute whatever values missing with a median value within each industry-month. This requires reasonable coverage by Compustat fundamentals and WRDS Financial Ratios datasets for each stock-month under consideration in addition to availability of conference call data[17].

Our baseline CATHlike portfolios follow the construction of the original S&P500 Catholic Values Index: the portfolio is long only; rebalancing is quarterly with constituents updated in

[17] A smaller number of eligible stocks before 2016, an issue that arises due to data availability, makes our ultimate task harder and our performance, therefore, is a conservative lower bound.

March, June, September and December; weights are value-weighted with a cap such that no individual stock has a weight greater than 24%; and the sum of weights of stocks with weight above 4.8% cannot be greater than 50%. For the sake of comparison, the only S&P500 stocks included into our CATHlike 500 portfolio, the portfolio closest to CATH, that are not in the original CATH in at least 10% of the months are Regeron Pharma (REGN-29% of months), Biogen (BIIB -24%), Eli Lilly (LLY -24%), Costco (COST -17%), Leidos Holdings (LDOS -16%), Paramount Global (PARA -14%) and Ball Corporation (BALL -10%).[18]

4. Optimization using Reinforcement Learning and Results

In previous sections we described the need to close the performance gap between values-based and unconstrained investing paradigms. We further showed how we can mimic and expand portfolio composition of ETFs such as CATH to years long before CATH inception and to stocks outside S&P500. In this section, we demonstrate how using advanced deep learning techniques we can close this gap through (1) portfolio optimization and (2) inclusion of stocks outside S&P500.

The economic environment is incredibly complex, ever-changing, and is represented by generally non-stationary observable variables of high dimensionality, all of which prove to be major challenges for any characteristics-based portfolio optimization technique. Being able to account for complex relationships among a large number of variables and to adapt dynamically to changes over time gives Reinforcement Learning methods an edge over traditional methods.

At the core of reinforcement learning is learning what actions to take upon observing a certain state of the environment. This mapping from states to actions is called a policy function. Having decided on the action, the environment transitions to the next state and rewards from actions taken at the previous step are revealed. The learner's task is to figure out by trial-and-error what actions in response to what state of the environment yield greatest cumulative rewards.

Mathematically speaking, let there be N_t eligible stocks at time t. Given time t state $s_t \in S_t \equiv \mathbb{R}^{N_t \times 12 \times K}$, a 12-month history for each stock of K stock-level predictors ($K = 52$ in our case), the agent is to choose an action $\boldsymbol{a_t} \in A_t \equiv [0,1]^{N_t}$, a vector of portfolio weights for each

[18] These firms are a varied lot ranging from aluminum packing (BALL) to entertainment (PARA), consumer products (COST), software (BIIB), and pharmaceuticals.

stock, from policy function $\pi: S \rightarrow A$. Reward at time t, r_t, is a function of both the time t and time $t+1$ states as well as the action taken at time t. It is customary to define a state-action-reward triplet $(s_t, \boldsymbol{a_t}, r_t) = (s_t, \boldsymbol{a_t} \in \pi(s_t), r_t)$ fully describing the game at time t. Given an episode of T such periods, the learner generates a state-action-reward trajectory $\tau = \{s_1, \boldsymbol{a_1}, r_1, \dots, s_T, \boldsymbol{a_T}, r_T\}$. In the context of our problem, $\boldsymbol{a_t}$ is a vector of holding weights between 0 and 1 of each stock in portfolio at time t sampled from the policy function π parametrized by a deep neural network. s_t are the input variables describing the state of the environment and r_t is the holding return of the corresponding portfolio. We set $T = 12$, that is for 12 consecutive months at a monthly frequency the agent rebalances their portfolio using the latest available information on the state of the environment and is rewarded with a monthly return. The parameters of the deep learning model, the policy function, are improved once the whole trajectory over 12 months is observed.

Our goal is to maximize the Sharpe ratio computed within each year from monthly returns across trajectories. We use return on an equal-weighted portfolio of the N_t eligible stocks as the benchmark in computing the Sharpe ratio instead of a risk-free rate in the objective function to maximize during training. The model learns from how well it does in terms of the objective function, the Sharpe ratio. Sometimes the economic environment sets a long-only trader up for failure such as during economic crises and there is little the model restricted to long-only positions in N_t stocks could have done to have a positive return that year. To avoid over penalizing the model in such instances, we reward the model for doing better than the equal-weighted index instead. Keep in mind that we use the conventional definition of the Sharpe ratio with the risk-free rate as the baseline in reporting any of our performance metrics.

Denoting parameters of the policy function after k epochs ("full" pass of the training data through the model) as θ_k, action at time t, $\boldsymbol{a_t}$, is sampled from policy $\pi_{\theta_k}(s_t)$ where the current state of the environment, s_t, also affects that action sampling distribution. We can also denote the vector of returns from trajectory τ generated using policy with parameters θ as r_τ^θ. The objective function of the agent is then $J(\theta) = Sharpe(r_\tau^\theta)$ and it is optimized to find the optimal weights of the policy function $\theta^* = argmax_\theta\, J(\theta)$ via gradient ascent as $\theta_{k+1} = \theta_k + \eta \nabla J(\theta)|_{\theta=\theta_k}$ at a learning rate η.

The model architecture closely resembles AlphaPortfolio by Cong et al (2025). The model has three components (see Figure 4). The first component is a sequence feature extractor which constructs a representation of each stock from its 12-month history of P attributes using a Transformer Encoder. The predictors are based on Compustat, CRSP and WRDS Financial Ratios. We list variable definitions in the appendix. In a sense, the sequence feature extractor re-expresses each predictor as a function of other predictors guided by how predictors are related to each other across time. The second component, a cross-attention network, captures interrelationships among stocks. It is based on a Multihead Attention module. Finally, winning scores are constructed using the third component by a linear layer transformation of the outputs from the cross-attention network and provided to the Reinforcement Learning agent who decides on action $\boldsymbol{a_t}$, a vector of portfolio weights. We focus on long-only portfolios and apply a Softmax operation to winning scores which rescales weights to be between 0 and 1 and to sum to 1 to determine portfolio weights, $\boldsymbol{a_t}$. The RL agent's goal is to maximize episode Sharpe ratio of the portfolio chosen by the agent.

We consider data from 1975 up to 2022 and split the sample into distinct training and test sets. More specifically, due to conference call data availability we use 1975-2003 years for training and 2004-2022 for testing. During training we split the training set into consecutive 2-year windows and each epoch, corresponding to one "full" pass of the data through the model, we sample a month within the first year of each 2-year window. The agent then interacts with the environment in the 12 consecutive months following that randomly chosen month. This data augmentation allows the agent to adapt to slightly different representations of similar environments reducing risks of overfitting and facilitating exploration. We repeat this training schedule for 30 epochs.

Each month the agent observes states of all stocks, both eligible and ineligible, and holds a long portfolio in 10% of stocks with the highest winning score irrespective of CATHlikeness or eligibility status. This way, the agent first becomes an expert in portfolio management learning from stocks both aligning with values-based investing and not. The agent's choice set will further be restricted to eligible stocks during testing. In Figure 5 we depict the in-sample learning curve. The red line representing the average annualized Sharpe ratio for epochs across trajectories is

roughly upward slowing and approaches a Sharpe ratio of 1.1 by end of training. All portfolios are long-only.

During testing, from 2004 onward, on the other hand, we sequentially apply the model to each new consecutive year restricting the choice set to only eligible or CATHlike stocks, record the agent's performance and then let the agent learn from its mistakes by taking 3 gradient ascent steps in the direction maximizing Sharpe ratio on this year. Importantly, since (1) conference call coverage is substantially better starting May of 2017; (2) CATH ETF started trading in 2016 and (3) we have holding data from ETF Global starting January 2017, we can only make meaningful comparisons between SPY, VTI, CATH ETF, optimized CATH and CATHlikes for the period limited to 2017-2022. We, therefore, use the out-of-sample period before 2017 starting in 2004 to transform our agent from an expert in unconstrained portfolio management to one proficient in values-based investing and thus ready for "prime time" which starts in year 2017.

We propose several ways to close the unfortunate performance gap between CATH and SPY. First, we ask if the underperformance of CATH, being limited to a subset of S&P500 stocks, is due to its having measurably fewer stocks included than SPY? We investigate whether simply adding stocks outside S&P500 that still look CATHlike and value weighting would boost performance of our synthetic values-based fund. Second, we ask whether value-weighting being a suboptimal portfolio optimization schedule is the cause of the underperformance? We apply our Reinforcement Learning algorithm to take stocks actually held by CATH but replace holding weights with optimal holding weights (we refer to this RL optimized portfolio as "CATH reweight"). Finally, we ask if using our methodology unconstrainted by portfolio size and value weighting can overcome the performance gap? To do so, we optimize portfolio holding weights of CATHlike portfolios.

In Table 3 we report annualized Fama-French 4 and 5 alphas, average returns, Sharpe ratios, Value-at-Risk, the corresponding Expected Shortfall at 5% and average number of stocks in each portfolio for VTI, SPY, CATH, CATH reweight, value-weighted and RL optimized CATHlikes for the period 2017-2022. A number of observations can be made.

First, CATH performed measurably worse than SPY in terms of all performance metrics, while also having more negative VAR and Expected Shortfall at 5%. Earlier we showed that CATH generally underperformed its benchmarks with respect to returns. But now we

demonstrate that CATH not only has worse returns, it also bears more risk than SPY. That said, CATH also had 50 stocks less than SPY in its portfolio on average throughout the period.

Secondly, our synthetic value-weighted CATHlikes perform comparably to SPY in terms of Sharpe ratios. We therefore show that adding largest n stocks outside S&P500 that we classify as CATHlike to the portfolio seems to close the gap in performance and is a much better alternative to simply dropping unCATHlike stocks within S&P500 and redistributing weights within industries.

Thirdly, our results highlight the limitations that value-weighting imposes on performance. We start by demonstrating that reweighting actual CATH holdings using our RL algorithm brings substantial improvements in performance metrics over the real CATH: Sharpe ratio of 0.67 vs 0.62; Fama-French 4 and 5 annualized alphas of 13.46%, 11.71% vs 0.36% and -0.1%. Notice that due to data availability the CATH reweight portfolio on average has 6.4% less stocks than actual CATH. The RL reweighted CATH is notably riskier than the true CATH with the standard deviation of monthly returns being 66% higher than CATH itself. The average returns, however, are 80% higher which more than compensates for the added risk.

Finally, RL optimizing CATHlike portfolios yields substantial gains and the greater the number of CATHlike stocks we include, the better the performance overall. For instance, the RL optimized CATHlike portfolio with 500 stocks has a Sharpe ratio of 0.92 compared to value-weighted CATHlike 500's 0.60, CATH ETF's 0.62 and SPY's 0.64. RL optimized CATHlikes also command a large Fama-French 4 and 5 annualized alpha, 15% and 19% respectively for RL optimized CATHlike 500 and 62% and 70% for RL optimized CATHlike 1500, implying that the performance cannot be attributed to a greater exposure to well-known risk factors. In Figure 6 we juxtapose cumulative returns since 2017 of value-weighted and RL optimized CATHlikes, CATH, SPY and RL reweighted CATH demonstrating substantial benefits from RL optimization and from forming larger portfolios.

One might ask why does reweighted CATH not perform nearly as well as the RL optimized CATHlike 500? The answer lies in the learning involved in the RL algorithm. In Figures 6 and 7 we plot cumulative returns of CATH reweighted and CATHlike 500 respectively against SPY, VTI and CATH. We see a substantial dip in the first year of the CATH reweighted portfolio (in 2017), a dip that is not observed for any other portfolio and definitely not for RL

optimized CATHlike 500. The agent in charge of CATHlike 500 had an opportunity to become a values-based investing expert, started implementing its strategies in 2004 and has been improving ever since. Unlike CATHlike 500, the agent making decisions on CATH reweighted holding weights just arrived at its post in 2017 after years of general portfolio management training in mostly the 20th century and, after an unsuccessful first year in the job, quickly learned how to do better. This means that, all else equal, it is reasonable to expect the CATH reweighting agent to continue improving and to eventually perform comparably to the CATHlike 500 manager. It also means that there likely is a better way to pretrain the CATH re-weighting manager using the data prior to 2017.

As stated above, Sharpe ratio improvements from RL over the period in question are substantial, but how stable are these improvements? In Table 4 we compare out-of-sample Sharpe ratios across portfolios by year. Several things are apparent from this table. The range of Sharpe ratios for value-weighted portfolios including CATH, SPY and VTI is substantially larger than that for the RL optimized portfolios with the lowest Sharpe ratio of -0.68 and largest of 4.34 among value-weighted portfolios and -0.23 and 2.99 respectively for RL optimized. Sharpe ratio of value-weighted portfolios is highly variable year-over-year unlike that of RL optimized portfolios with the average standard deviation of Sharpe ratios among value-weighted portfolios being 1.67 and that for RL optimized portfolios being 0.8. This difference is statistically significant at 1%. This implies that RL optimized portfolios provide more consistent returns per unit of risk than value-weighted portfolios. Most importantly, CATHLike 500, the portfolio we are most interested in since it essentially is an RL optimized CATHlike version of SPY, never experienced a negative Sharpe ratio in the out-of-sample period considered unlike all other portfolios and ETFs in the comparison group and stayed around 1 in most years.

5. Rationalizing RL Performance: Turnover, Transaction Costs and Stock Characteristics

5.1 Turnover and Transaction Costs

One common belief is that the impressive performance of many machine learning models in finance is attributable to trading frictions such as illiquidity of stocks traded by the model. In this subsection we address these concerns by examining the robustness of our results to trading costs. Such costs are generally much higher for small and low liquidity stocks so including these costs

allows us to capture at least partially this illiquidity dimension. We demonstrate that our model performs well even after accounting for trading costs.

We start by constructing a measure of portfolio turnover following Gu et al (2020). More specifically, average monthly turnover of a portfolio is defined as

$$Turnover = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\left|\frac{w_{i,t}(1+r_{i,t+1})}{1+\sum_j w_{j,t}r_{j,t+1}} - w_{i,t+1}\right|$$

with $w_{i,t}$ being portfolio weight of stock i at time t and $r_{i,t}$ the corresponding monthly return. Gu et al (2020) demonstrate that machine learning models often outperform at the cost of a greater portfolio turnover and virtually all machine learning models considered in that paper have a turnover of at least 110%.

In Table 5 we report turnovers for our RL optimized and value-weighted portfolios. Value-weighted portfolios have a dramatically smaller turnover - less than 10% for all portfolio sizes. We expect SPY and CATH to have a comparable turnover estimate as these are constructed in a similar manner to our value-weighted synthetic CATHlikes. Turnover of RL optimized portfolios, however, largely depends on portfolio size. Nevertheless, nearly all RL optimized portfolios have turnovers smaller than 110%, the minimum turnover reported in Gu et al (2020), with RL reweighted CATH, RL optimized CATHlike 500 and 750 having a turnover of 63%, 83% and 109% respectively. Only portfolios with 1000 and 1500 stocks have a higher turnover – 122% and 136% respectively.

Finally, we follow Bollerslev et al (2018) and account for transaction costs using the median bid-ask spread in the last 9 months for each stock. We adjust for transaction costs equal to a double "half-spread" or one full bid-ask spread as is done in that paper, but also consider costs equal to 2 full spreads and 5 full spreads. We report performance metrics adjusted for bid-ask spreads in Table 6.

As expected, performance is decreasing in transaction costs. Most importantly, the performance of RL optimized models holds up remarkably well. Even if adjusted for 5 full bid-ask spreads, RL optimized CATHlike 500, 750, 1000, 1500 and CATH reweight portfolios boast a Sharpe ratio of 0.82, 0.81, 1.03, 1.17 and 0.64 respectively, in comparison to value-weighted CATHlike 500's Sharpe ratio of 0.6. Fama-French 4 and 5 alphas of our RL optimized portfolios

stay substantial and significant throughout as well. The performance of our model is thus robust to transaction costs.

5.2 Stock Characteristics

How does the RL model achieve this performance? The demand for explainability of AI models in social sciences is high. As an attempt to explain our model's performance, we further investigate what stock characteristics are associated with a larger or smaller optimal portfolio weight allocation as suggested by our RL agent. More specifically, we start by computing differences in weights between each RL optimized portfolio and the actual CATH. We use industry definitions and stock characteristics, the information set originally available to the RL agent in making decisions, as the explanatory variable set. We carry out a quantile transformation within each month of each non-binary variable and regress the weight difference on this set of explanatory variables, but first we employ a lasso regression to reduce the number of predictors and limit multicollinearity in the explanatory variable set.

Regression equation is thus

$$\left(\hat{\alpha}, \widehat{\boldsymbol{\beta}}\right) = argmin_{\alpha\in\mathbb{R},\boldsymbol{\beta}\in\mathbb{R}} \|\boldsymbol{w} - \alpha\mathbf{1} - \boldsymbol{X}\boldsymbol{\beta}\|_2^2 \quad + \lambda\|\boldsymbol{\beta}\|_1$$

where $\boldsymbol{w}$ is the weight difference vector that we are trying to explain, $\boldsymbol{X}$ is a matrix containing predictors by stock and over time, α is the intercept, $\boldsymbol{\beta}$ is the vector of coefficients of interest indicating attribution of each stock characteristic and λ is a hyperparameter controlling the strength of Lasso regularization.

The results are reported in Table 7 and all variable definitions can be found in the Appendix. First, notice that the signs of each coefficient and their statistical significance are uniform across portfolios for most explanatory variables. This means that whatever the size of the portfolio, the RL agent eventually arrives at a somewhat similar decision rule in stock selection. Stocks in the medical equipment, healthcare, machinery and transportation industries as well as those with a higher market capitalization received a smaller weight than in actual CATH. A higher assets-to-market-capitalization ratio, price-to-sales ratio, Tobin's Q, cashflow relative to total liabilities, SG&A to sales, turnover (trading volume over shares outstanding), cash holdings to assets, net operating assets, dividend price ratio, absolute operation accruals, new share issue, investment, gross margin growth relative to sales growth as well as stocks in insurance, restaurants, hotel

business received a higher weight than in actual CATH. The model thus recovers the size and value factors, favors stocks well positioned to stay profitable, those with a good investment opportunity set and with recent information regarding the stock.

5. Conclusions

Incorporating values-based restrictions into portfolio selection has long been viewed as inducing a "doing good versus doing well" tradeoff for investors. Using the tools of artificial intelligence, we have demonstrated that this outcome need not be inevitable. Drawing on recent advances in Large Language Models and Deep Reinforcement Learning, we showed how risk-adjusted returns can be substantially improved via enhanced asset selection and more complex portfolio weighting optimization. Our application is to faith-based investing, where issues related to values-appropriate asset selection are particularly challenging, and shorter trade histories combined with value-weighted investment vehicles often limit investment returns. This faith-based setting seems particularly well-suited to AI-based analytics.

Can this approach be applied more generally? We believe our AI-based approach is applicable to any portfolio optimization problem that is subject to textual, nebulous, or implicit constraints. As noted earlier, broader SRI mandates can fall into this category, but so, too, can restrictions on portfolios arising from tax mandates, or geographic limitations, or simply thematic preferences such as sentiment. An interesting question we want to address in future research is the impact on returns and risk imposed by diversification-based regulatory constraints on portfolio weights and construction. In our analysis here, we stayed in within the confines dictated by such constraints as they applied to CATH but still found that we could create portfolios with lower sharp ratios. We conjecture that such regulatory constraints, while useful in times past for restricting risk, may now be more of an impediment than a protection for investors risk exposure.

One aspect of our research that we think important to emphasize is the role played by advances in artificial intelligence in explaining our out-performance. AI-based investing is not actually new, with the first AI-based ETF (AIEQ) introduced in late 2017. AIEQ used IBM Watson to select stocks, but the resulting ETF had a lower Sharpe ratio and higher risk than its

benchmark. As a result, it has underperformed the S&P index every year since its inception.[19] Newer generations of AI-based ETFs have emerged, but again their track record has been mixed, with some posting particularly poor performance.[20] The advent of transformer-based architecture has revolutionized AI analyses, paving the way for greater success in crafting investment strategies. The success of our methodology, which relies on machines rather than a person making portfolio decisions, suggests that such algorithmic-based investing can play an important role in asset management.

[19] The current year is no exception. 2024 year to date returns are 5.59 compared to SPY returns of 22.36.
[20] An interesting discussion of AI-based ETF performance is "Hype vs Reality: How are AI-Driven ETFs Performing?" at https://www.dividend.com/active-etfs-channel/how-are-ai-driven-etfs-performing/

References

Berk, J. B. and van Binsbergen, J. H., 2025, The Impact of Impact Investing, Journal of Financial Economics, available at https://doi.org/10.1016/j.jfineco.2024.103972.

Bollerslev, T., Benjamin Hood, John Huss, Lasse Heje Pedersen, Risk Everywhere: Modeling and Managing Volatility, The Review of Financial Studies, Volume 31, Issue 7, July 2018, Pages 2729–2773, https://doi.org/10.1093/rfs/hhy041

Bolton, P., and M. Kacperczyk, 2021, Do Investors Care about Carbon Risk? Journal of Financial Economics, 142(2) 517-549.

Cao S., Wei Jiang, Baozhong Yang, Alan L Zhang, How to Talk When a Machine Is Listening: Corporate Disclosure in the Age of AI, The Review of Financial Studies, Volume 36, Issue 9, September 2023, Pages 3603–3642, https://doi.org/10.1093/rfs/hhad021

Cong, W., K. Tang, J. Wang, and Y. Zhang, 2025, Alpha Portfolio: Direct Construction Through Deep Reinforcement Learning and Interpretable AI, Journal of Financial Economics, forthcoming.

Chava, S., W. Du, and B. Malakar, 2021, Do Managers Walk the Talk on Environmental and Social Issues, Available at SSRN: https://ssrn.com/abstract=3900814

De Nard, G., Engle, R.F. and B. Kelly, 2024, Factor Mimicking Portfolios for Climate Risk, Financial Analysts Journal, 80:3, 37-58, DOI: 10.1080/0015198X.2024.2332164Edmans, A. and M. Kapercyzk, 2022, Sustainable Finance, Review of Finance, 26(5), 1309-1311.

Gabaix, Xavier and Koijen, Ralph S. J. and Richmond, Robert and Yogo, Motohiro, Asset Embeddings (September 14, 2023). Available at SSRN: https://ssrn.com/abstract=4507511 or http://dx.doi.org/10.2139/ssrn.4507511

Green, D. and B. Roth, 2054, The Allocation of Socially Responsible Investing, Journal of Finance, 80(2) 755-781.

S. Gu, B. Kelly, D. Xiu, "Empirical Asset Pricing via Machine Learning," 2020, Review of Financial Studies, 33, 2223-2273

Hong, H. and M. Kacperczyk, 2009, The price of sin: the effects of social norms on markets, Journal of Financial Economics, 93(1) 15-36.

Kelly, Bryan T. and Xiu, Dacheng, 2023, Financial Machine Learning , Foundations and Trends in Finance.

Kelly, Bryan T. and Malamud, Semyon and Zhou, Kangying, 2024. The Virtue of Complexity in Return Prediction, Journal of Finance, 79(1) 459-503.

Koijen, Ralph S. J. Yogo, Motohiro, 2019, A Demand System Approach to Asset Pricing, Journal of Political Economy, 127(4). 1475-1515.

C. Krauss, X. A. Do and N. Huck, "Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500", European Journal of Operational Research, Volume 259, Issue 2, 1 June 2017, Pages 689-702

Krueger, P., Z. Saunter, and L. Starks, 2020, The importance of climate risks for institutional investors, Review of Financial Studies, 33, 1067-1111.

Loughran, T. and McDonald, B. (2011), When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. The Journal of Finance, 66: 35-65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Matvos, P., 2020, ESG and Responsible Investing around the World: A Survey, CFA Institute

Morningstar Research, "A matter of faith", Dec. 6, 2024, 2-23.

Murfin, J. and M. Spiegel, 2020, Is the risk of sea level rise capitalized in residential real estate? Rev. Financ. Stud., 33 (2020), pp. 1217-1255

Pastor, L., R. Stambaugh, L. Taylor, 2021, Sustainable investing in equilibrium, Journal of Financial Economics, 142(2) 550 -571.

Pastor, Lubos and Stambaugh, Robert F. and Taylor, Lucian A., Dissecting Green Returns (June 10, 2022). University of Chicago, Becker Friedman Institute for Economics Working Paper No.2021-70,Available at SSRN: https://ssrn.com/abstract=3869822 or http://dx.doi.org/10.2139/ssrn.3869822

Petersen, L. H., S. Fitzgibbons, and L. Pomorski, 2021, Responsible Investing: The ESG-efficient Frontier, Journal of Financial Economics, 124(2) 572- 601.

Starks, L, 2023, Presidential Address: Sustainable Finance and ESG Issues – Value versus Values, Journal of Finance, 78(4) 1837-1872.

Yang, Keer and Yasuda, Ayako, 2023, Decoding Sustainable Investment Strategies: Bridging Intentions and Actions, available at SSRN: https://ssrn.com/abstract=4637264 or http://dx.doi.org/10.2139/ssrn.463726

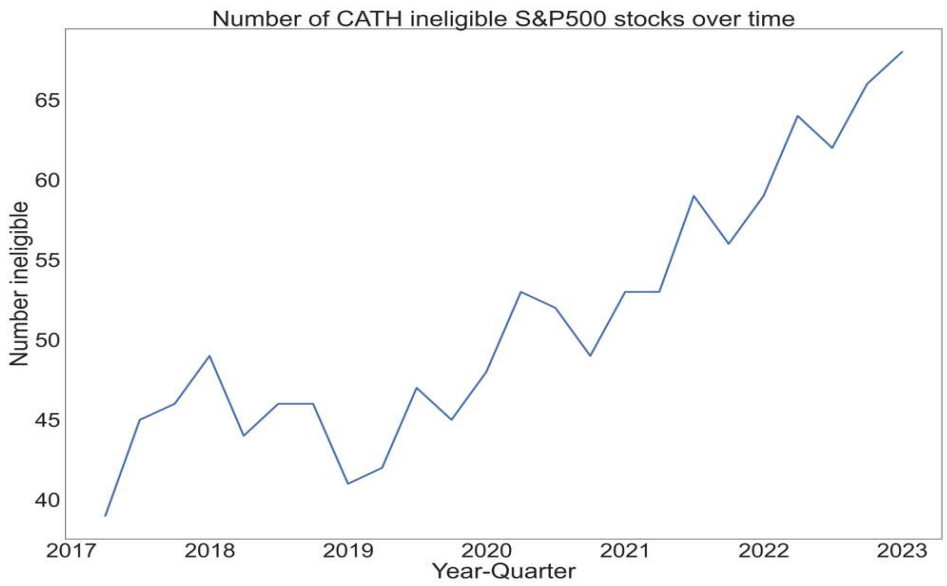


Figure 1. The number of CATH ineligible stocks over time

Note. Here we plot the number of stocks in the S&P500, but not in CATH over time at a quarterly frequency. These excluded stocks are CATH ineligible as classified by the CATH ETF management.

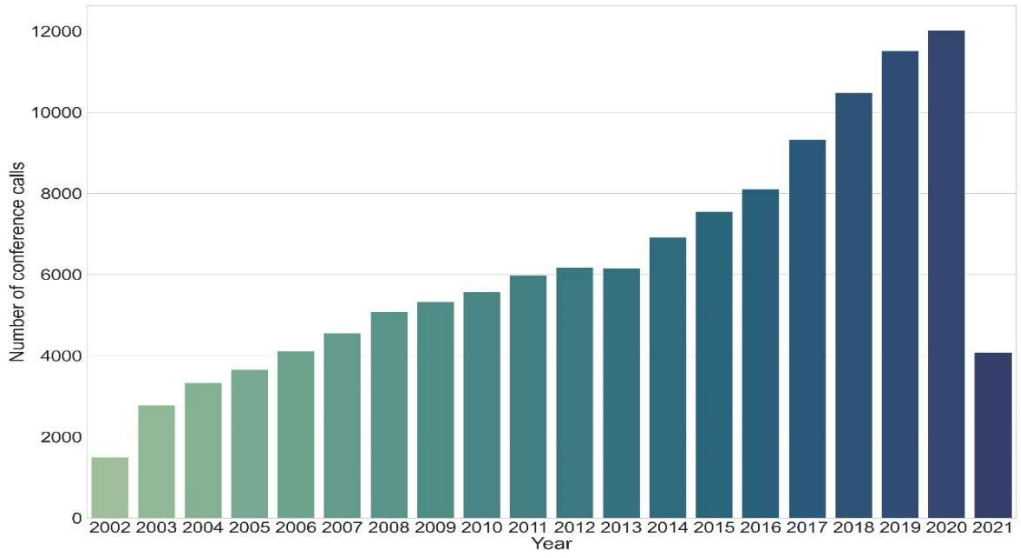


Figure 2. Number of Conference Calls over time

Note. This Figures plots the number of conference calls in our sample over time. The data provider is Refinitiv and coverage is increasing over time and ends mid-2021.

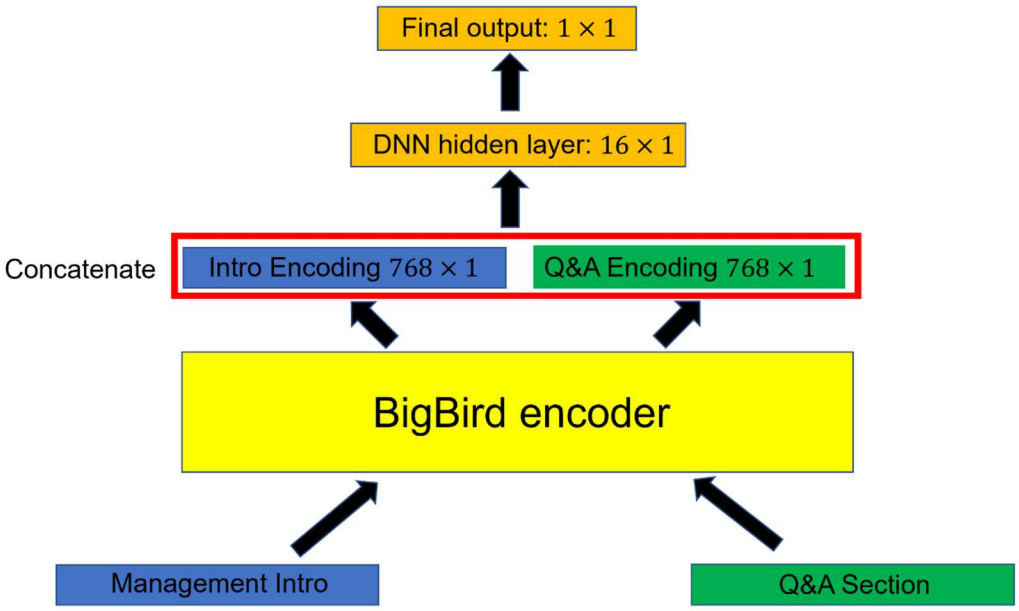


Figure 3. Large Language Model Architecture

Note. This Figures depicts the flowchart representing our LLM model. The model processes the two conference call sections, management intro and Q&A, in parallel constructing a 768-dimensional representation of each. Together this conference call representation is fed into a 1-hidden layer neural network predicting market reaction to conference call. The 16-dimensional features learned by the deep neural network in predicting market reaction are the features we use for the purposes of detecting CATH-ineligible stocks.

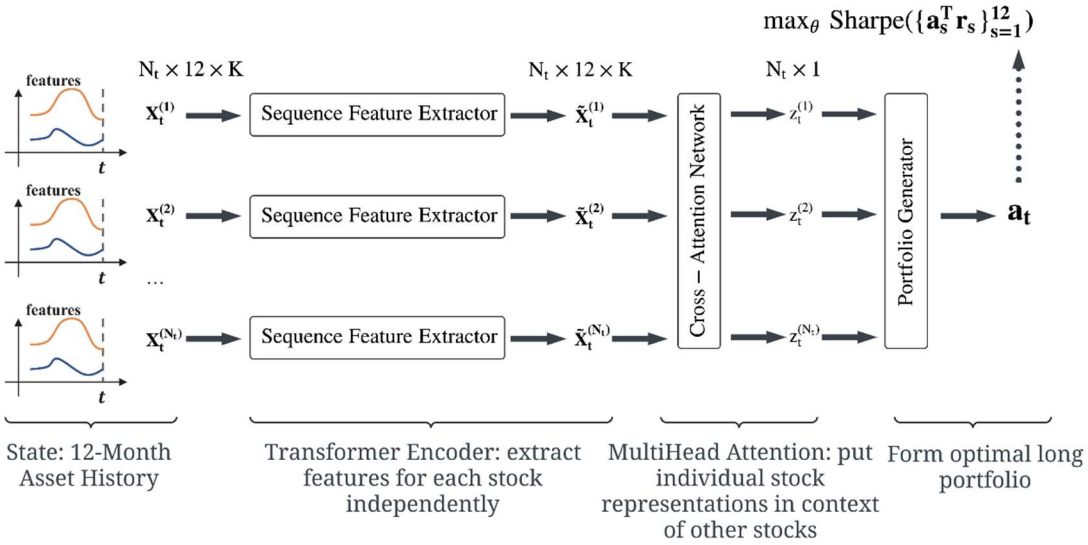


Figure 4. Proposed Reinforcement Learning pipeline

Note. This graphic depicts our Reinforcement Learning pipeline used to optimize portfolio weights. We observe a 12-month history of firm-level observations for each stock. The model first extracts features independently for each stock across the time dimension. These features are then put in context of other stocks via the cross-attention mechanism. The agent finally learns to construct winning scores based on these relative features and to construct portfolios.

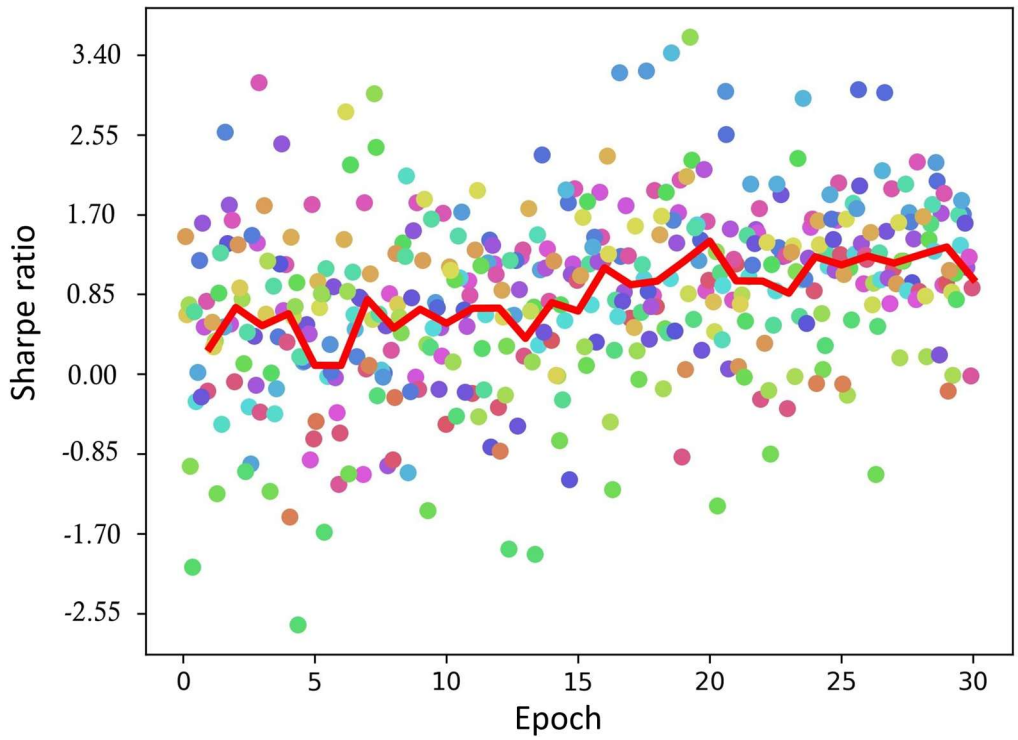


Figure 5. In-sample Learning Curve of our Reinforcement Learning Algorithm

Note. This plot depicts in-sample improvement of the algorithm over time in terms of portfolio Sharpe ratio. The red line is the average Sharpe ratio per episode across the sampled years with each individual dot representing that year's Sharpe ratio performance.

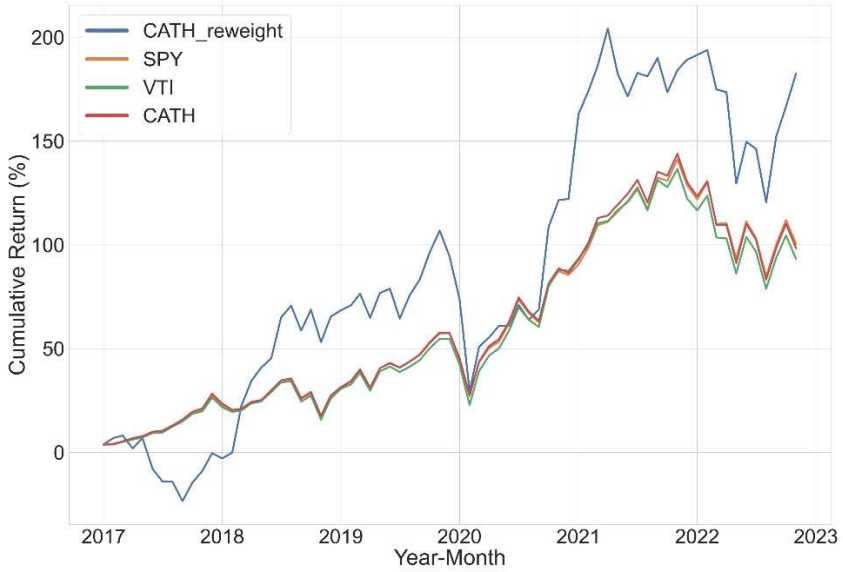


Figure 6. Cumulative returns of CATH reweighted, CATH, SPY and VTI

Note. This Figure compares out-of-sample cumulative return of baseline ETFs and portfolios such as SPY, VTI and CATH and an RL reweighted CATH portfolio between 2017 and 2023.

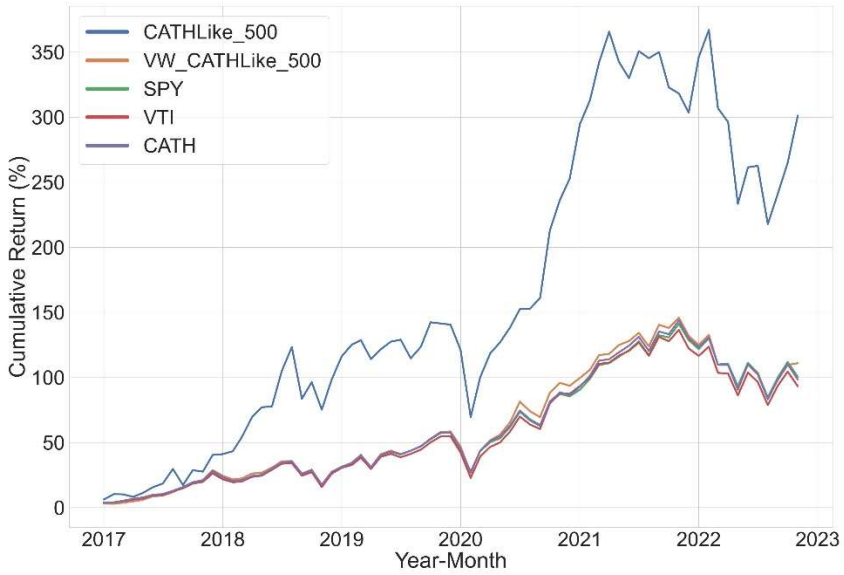


Figure 7. Cumulative returns of RL optimized and value-weighted CATHlike 500, CATH, SPY and VTI

Note. This plot compares out-of-sample cumulative return between 2017 and 2023 of baseline ETFs and portfolios such as SPY, VTI and CATH, the value-weighted CATHLike portfolio with 500 stocks and an RL reweighted CATH portfolio.

	Baseline (naïve)	FF48 dummies	Improvement over Baseline (%)	Model	Improvement over Baseline (%)
Accuracy	0.88	0.79	-10	0.93	6
Precision	0.12	0.33		0.82	
Recall	1	0.64		0.62	
F1-score	0.21	0.43	105	0.7	233
Average Precision	0.12	0.35	192	0.75	525
ROC-AUC	0.5	0.79	58	0.88	76

Table 1. CATH ineligibility classification: out-of-sample performance

Note. In this table we compare out-of-sample classification performance of our model in detecting CATH ineligible stocks with the naïve baseline and an alternative logistic regression model based on industry definitions alone. The comparisons are carried out along several metrics: accuracy, precision, recall, F1-score, Average Precision and ROC-AUC score with Average Precision being the most appropriate for this setting given that this is an anomaly detection problem with a major class imbalance. A higher value of each measure implies a better model.

	Q1	Q2	Q3	Q4	Q5	Q1-Q5
Average return (%)	0.89***	0.86**	0.74**	0.71**	0.64**	0.25**
	3.17	2.39	2.39	2.31	2.07	2.47
FF4 alpha (%)	2.87***	1.07	-0.05	-0.2	-0.7	3.59***
	2.74	0.87	-0.05	-0.16	-1.05	2.99

Table 2. Comparison of average returns and annualized Fama-French 4 alphas across CATHlike sorts since 2004

Note. In this table we juxtapose average returns and Fama-French 4 alphas for CATHlike portfolio sorts from 2004 onward. Our Natural Language Processing pipeline uses conference calls to produce a probability for each quarter that a stock is ineligible for inclusion into CATH. Within each quarter we sort stocks by their probability of inclusion into CATH or their "CATHlikeness". Probabilities are fixed within each quarter and we use the latest conference call held before each quarter for sorts in each respective quarter. We report t-statistics under each estimate derived from the Newey-West estimator with 5 lags. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

	Annualized alpha (%)		Average return	Sharpe	VAR 5%	Exp Shortfall 5%	Number of stocks
	FF4	FF5					
CATHLike 500	15.05**	18.76***	2.11	0.92	-12.81	-17.67	500
CATHLike 750	22.6*	29.15**	2.58	0.89	-14.19	-19.95	750
CATHLike 1000	37.17***	40.65***	3.22	1.13	-13.71	-19.05	1000
CATHLike 1500	62.16***	69.8***	4.33	1.28	-12.77	-19.43	1500
CATH reweight	13.46	11.71	1.64	0.67	-11	-16.84	426
VW CATHLike 500	1.04	1.15	1.01	0.67	-9.4	-10.93	500
VW CATHLike 750	0.91	1.06	0.99	0.65	-9.48	-11.05	750
VW CATHLike 1000	0.92	1.04	0.99	0.65	-9.51	-11.13	1000
VW CATHLike 1500	0.92	1.03	0.99	0.64	-9.49	-11.22	1500
SPY	0.57	0	0.92	0.64	-8.52	-9.92	505
VTI	0.02	-0.18	0.88	0.59	-8.73	-10.47	3567
CATH	0.36	-0.1	0.91	0.62	-8.99	-10.23	455

Table 3. Comparison of out-of-sample performance of value-weighted and RL optimized CATHLikes, RL reweighted CATH with true CATH, SPY and VTI over 2017-2022 period

Note. In this table we juxtapose out-of-sample performance of Reinforcement Learning (RL) optimized and unoptimized portfolios from 2017 to 2022, the years when CATH was actively traded. SPY, VTI and CATH are used as baselines. CATHLike portfolios are constructed using our Natural Language Processing model classifying stocks as CATH-ineligible from text. We construct synthetic value-weighted CATHLike portfolios of different sizes with 500, 750, 1000 and 1500 stocks in each. We further optimize these portfolios using RL. We thus compare baseline ETFs such as SPY, VTI and CATH with the value-weighted CATHLike synthetic ETFs as well as with RL optimized CATH and CATHLike portfolios. As basis for comparison, we use Fama-French 4 and 5 alphas, Sharpe ratios, Value-at-Risk (VAR) and Expected Shortfall (SF) at 5% as well as the average number of stocks in each portfolio. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

Portfolio/Year	2017	2018	2019	2020	2021	2022
CATHLike 500	1.75	1.09	1	1.16	0.8	0.05
CATHLike 750	2.03	1.45	0.3	0.85	0.46	-0.1
CATHLike 1000	2.35	1.21	1.36	1.3	0.62	-0.07
CATHLike 1500	2.99	1.57	0.8	1.79	0.26	-0.23
CATH reweight	0.05	1.73	0.73	0.5	1.26	-0.04
VW CATHLike 500	4.08	-0.23	1.53	0.82	1.45	-0.44
VW CATHLike 750	4.08	-0.23	1.47	0.81	1.41	-0.42
VW CATHLike 1000	4.12	-0.23	1.44	0.81	1.4	-0.41
VW CATHLike 1500	4.14	-0.23	1.39	0.81	1.39	-0.4
SPY	4.07	-0.18	1.59	0.72	1.69	-0.63
VTI	4.13	-0.16	1.48	0.8	1.36	-0.66
CATH	4.34	-0.19	1.55	0.74	1.65	-0.68

Table 4. Out-of-sample Sharpe ratios by year and portfolio

Note. In this table we compare out-of-sample Sharpe ratios by year of our Reinforcement Learning (RL) optimized and unoptimized portfolios from 2017 to 2022, the years when CATH was actively traded. SPY, VTI and CATH are used as baselines. CATHLike portfolios are constructed using our Natural Language Processing model classifying stocks as CATH-ineligible from text. We construct synthetic value-weighted CATHLike portfolios of different sizes with 500, 750, 1000 and 1500 stocks in each. We further optimize these portfolios using RL. We thus compare baseline ETFs such as SPY, VTI and CATH with the value-weighted CATHLike synthetic ETFs as well as with RL optimized CATH and CATHLike portfolios.

Portfolio	Turnover (%)
CATHLike 500	83.38
CATHLike 750	109.28
CATHLike 1000	121.67
CATHLike 1500	135.61
CATH reweight	62.53
VW CATHLike 500	9.33
VW CATHLike 750	9.54
VW CATHLike 1000	9.71
VW CATHLike 1500	9.91

Table 5. Portfolio turnover for value-weighted and RL optimized CATHLikes, RL reweighted CATH over 2017-2022 period following Gu et al (2020)

Note. In this table we juxtapose out-of-sample turnovers of Reinforcement Learning (RL) optimized and unoptimized portfolios from 2017 to 2022, the years when CATH was actively traded. CATHLike portfolios are constructed using our Natural Language Processing model classifying stocks as CATH-ineligible from text. We construct synthetic value-weighted CATHLike portfolios of different sizes with 500, 750, 1000 and 1500 stocks in each. We further optimize these portfolios using RL. We thus compare baseline ETFs such as SPY, VTI and CATH with the value-weighted CATHLike synthetic ETFs as well as with RL optimized CATH and CATHLike portfolios. Turnover measure is computed following Gu et al (2020).

	Annualized alpha (%)		Average return	Sharpe	VAR 5%	Exp Shortfall 5%
	FF4	FF5				
CATHLike 500 less ba x 1	12.85**	16.66***	1.95	0.87	-11.95	-17.65
CATHLike 500 less ba x 2	12.53**	16.34***	1.93	0.86	-11.96	-17.68
CATHLike 500 less ba x 5	11.59**	15.36**	1.86	0.83	-12.00	-17.75
CATHLike 750 less ba x 1	21.03*	27.68**	2.47	0.86	-14.21	-19.99
CATHLike 750 less ba x 2	20.52*	27.15**	2.43	0.84	-14.23	-20.03
CATHLike 750 less ba x 5	18.99*	25.59**	2.33	0.81	-14.29	-20.16
CATHLike 1000 less ba x 1	35.09***	38.66***	3.10	1.10	-11.63	-18.02
CATHLike 1000 less ba x 2	34.29***	37.85***	3.05	1.09	-11.69	-18.08
CATHLike 1000 less ba x 5	31.9***	35.44***	2.90	1.03	-11.85	-18.27
CATHLike 1500 less ba x 1	60.24***	68.55***	4.28	1.27	-12.84	-19.46
CATHLike 1500 less ba x 2	58.71***	66.92***	4.20	1.24	-12.90	-19.55
CATHLike 1500 less ba x 5	54.18***	62.11***	3.95	1.17	-13.11	-19.82
CATH reweight less ba x 1	12.93	11.6	1.62	0.67	-11.09	-16.37
CATH reweight less ba x 2	12.69	11.37	1.60	0.67	-11.10	-16.38
CATH reweight less ba x 5	12	10.69	1.55	0.64	-11.12	-16.42
VW CATHLike 500 less ba x 1	-0.11	0	0.91	0.60	-9.53	-11.06
VW CATHLike 500 less ba x 2	-0.12	-0.01	0.91	0.60	-9.53	-11.06
VW CATHLike 500 less ba x 5	-0.16	-0.05	0.91	0.60	-9.53	-11.06
VW CATHLike 750 less ba x 1	-0.23	-0.09	0.90	0.59	-9.59	-11.19
VW CATHLike 750 less ba x 2	-0.24	-0.11	0.90	0.59	-9.59	-11.19
VW CATHLike 750 less ba x 5	-0.28	-0.15	0.89	0.59	-9.59	-11.19
VW CATHLike 1000 less ba x 1	-0.23	-0.11	0.90	0.58	-9.55	-11.26
VW CATHLike 1000 less ba x 2	-0.24	-0.13	0.89	0.58	-9.55	-11.26
VW CATHLike 1000 less ba x 5	-0.28	-0.17	0.89	0.58	-9.55	-11.26
VW CATHLike 1500 less ba x 1	-0.23	-0.12	0.89	0.58	-9.53	-11.35
VW CATHLike 1500 less ba x 2	-0.24	-0.14	0.89	0.58	-9.53	-11.35
VW CATHLike 1500 less ba x 5	-0.29	-0.19	0.89	0.58	-9.53	-11.35

Table 6. Performance of value-weighted and RL optimized CATHLikes and RL reweighted CATH adjusted for bid-ask spreads (x1, x2, x5) over 2017-2022 period

Note. This table juxtaposes out-of-sample performance of Reinforcement Learning (RL) optimized and unoptimized portfolios from 2017 to 2022. We adjust performance of each portfolio for bid-ask spread multiples (× 1, × 2 and × 5). CATHLike portfolios are constructed using our Natural Language Processing model classifying stocks as CATH-ineligible from text. We construct synthetic value-weighted CATHLike portfolios of different sizes with 500, 750, 1000 and 1500 stocks in each. We further optimize these portfolios using RL. We thus compare value-weighted CATHLike synthetic ETFs with RL optimized CATH and CATHLike portfolios adjusted for transaction costs. As basis for comparison, we use Fama-French 4 and 5 alphas, Sharpe ratios, Value-at-Risk (VAR) and Expected Shortfall (SF) at 5%. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

Predictor\Portfolio	500	750	1000	1500	Reweight
ME	-0.63***	-0.75***	-0.79***	-0.82***	-0.51***
Medical Equipment	-0.16***	-0.17***	-0.18***	-0.19***	-0.17***
Healthcare	-0.12*	-0.09	-0.09	-0.09	-0.14***
Machinery	-0.05*	-0.04**	-0.04**		-0.06***
Transportation	-0.04*		-0.04**		
EPS	-0.03	-0.05**	-0.04**	-0.06***	-0.05**
pretret_noa	-0.03	-0.06***	-0.05**		-0.07***
AOA	0.02***	0.02***	0.03***	0.02***	0.01
Investment	0.02*	0.02**	0.02**	0.02*	0.02***
gmsale	0.03***	0.02***	0.03***	0.02***	0.01***
cshrout	0.03***	0.02***	0.02***	0.01**	0.03***
cso	0.04***	0.04***	0.04***	0.04***	0.02***
LDP	0.04	0.06**	0.06***		0.08***
NOA	0.05*	0.08**	0.07***	0.08***	
Restaurants, Hotels	0.05*		0.06***		
C	0.13***	0.28***	0.28***	0.29***	0.22***
Turnover	0.13***	0.09***	0.1***	0.06**	0.22***
SGA2S	0.15***	0.13***	0.11***	0.11***	0.1***
Insurance	0.17***	0.14***	0.14***	0.11***	0.11***
C2D	0.17***	0.13***		0.12***	0.16***
ps	0.18***	0.23***	0.26***	0.32***	0.1***
Q	0.18***	0.15***		0.11***	0.18***
A2ME	0.52***	0.4***	0.43***	0.44***	0.57***
S2C		0.21***	0.2***	0.24***	0.15***
Std_vol		-0.06***			-0.07***
Controls	Yes	Yes	Yes	Yes	Yes
Adjusted R2 (%)	94	95	92	93	93

Table 7. Explaining RL portfolio over- or underweighting of stocks

Note. This table investigates the decisions made by our RL agent. For each stock and each RL portfolio the dependent variables is the portfolio weight difference from actual CATH. We use industry definitions and stock characteristics, the information set originally available to the RL agent in making decisions, as the explanatory variable set. We carry out a quantile transformation within each month of each non-binary variable and regress the weight difference on this set of explanatory variables for each RL portfolio separately, but first we employ a lasso regression to reduce the number of predictors and limit multicollinearity in the explanatory variable set. All variable definitions are given in the appendix. Controls cover all other explanatory variables that were not statistically significant. Standard errors are clustered at the stock-level. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

Regression equation is

$$(\hat{\alpha}, \widehat{\boldsymbol{\beta}}) = argmin_{\alpha \in \mathbb{R}, \boldsymbol{\beta} \in \mathbb{R}} \|\boldsymbol{w} - \alpha \mathbf{1} - \boldsymbol{X}\boldsymbol{\beta}\|_2^2 \quad + \lambda \|\boldsymbol{\beta}\|_1$$

where $\boldsymbol{w}$ is the weight difference vector that we are trying to explain, $\boldsymbol{X}$ is a matrix containing predictors by stock and over time, α is the intercept, $\boldsymbol{\beta}$ is the vector of coefficients of interest indicating attribution of each stock characteristic and λ is a hyperparameter controlling the strength of Lasso regularization.

Appendix. RL Optimization Variable Definitions

Below we list variable definitions for the 52 predictors used in the RL optimization part of the paper. The data come from the Merged CRSP-Compustat, CRSP and WRDS Financial Ratio Firm Level databases. Most of the variable definitions follow Cong et al (2019).

A2ME: We define assets-to-market cap as total assets over market capitalization.

$$A2ME_t = \frac{AT_t}{SHROUT_t * PRC_t}$$

OA: We define operating accruals as change in non-cash working capital minus depreciation scaled by lagged total asset.

$$OA_t = \frac{\Delta\left((ACT_t + CHE_t - LCT_t - DLC_t - TXP_t) - DP_t\right)}{AT_{t-1}}$$

AOA: We define AOA as absolute value of operation accruals.

AT: Total asset.

BEME (WRDS Financial Ratios Firm Level database): Ratio of book value of equity to market equity.

Beta_daily: Beta daily is the sum of the regression coefficients of daily excess returns on the market excess return and one lag of the market excess return.

C: Ratio of cash and short-term investments to total assets.

$$C_t = CHE_t / AT_t$$

C2D: Cash flow to price is the ratio of income and extraordinary items and depreciation and amortization to total liabilities.

$$C2D_t = (IB_t + DP_t) / LT_t$$

CTO: We define capital turnover as ratio of net sales to lagged total assets.

$$CTO_t = SALE_t / AT_{t-1}$$

CAPEI: Shiller's Cyclically Adjusted P/E Ratio

Dept2P: Debt to price is the ratio of long-term debt and debt in current liability to the market capitalization.

$$\text{Dept2P}_t = \frac{DLTT_t + DLC_t}{SHROUT_t * PRC_t}$$

Δceq: The percentage change in the book value of equity.

$$\Delta ceq_t = (CEQ_t - CEQ_{t-1}) / CEQ_{t-1}$$

Δ(ΔGm - ΔSales:) The difference in the percentage change in gross margin and the percentage change in sales.

$$\Delta(\Delta Gm - \Delta Sales)_t = \frac{SALE_t - COGS_t}{SALE_{t-1} - COGS_{t-1}} - \frac{SALE_t}{SALE_{t-1}}$$

ΔSo: Log change in the split adjusted shares outstanding.

$$\Delta So_t = log(CSHO_t * AJEX_t) - log(CSHO_{t-1} * AJEX_{t-1})$$

Δshrout: Percentage change in shares outstanding.

$$\Delta shrout_t = (SHROUT_t - SHROUT_{t-1}) / SHROUT_{t-1}$$

ΔPI2A: The change in property, plants, and equipment over lagged total assets.

$$\Delta P12A_t = \frac{\Delta(PPENT_t + INVT_t)}{AT_{t-1}}$$

E2P: We define earnings to price as the ratio of income before extraordinary items to the market capitalization.

$$E2P_t = IB_t / (SHROUT_t * PRC_t)$$

EPS: We define earnings per share as the ratio of income before extraordinary items to shares outstanding.

$$EPS_t = IB_t / SHROUT_t$$

Free CF: Cash flow to book value of equity.

$$FreeCF_t = \frac{NI_t + DP_t + WCAPCH_t + CAPX_t}{BE_t}$$

Idol vol: Idiosyncratic volatility is the standard deviation of the residuals from a regress of excess returns on the Fama and French three-factor model.

Investment: We define investment as the percentage year-on-year growth rate in total assets.

$$Investment_t = (AT_t - AT_{t-1}) / AT_{t-1}$$

IPM (WRDS Financial Ratios Firm Level database): Pretax profit margin, *EBT/Revenue*.

IVC: We define IVC as change in inventories over the average total assets of t and $t-1$

$$IVC_t = \frac{2 * (INVT_t - INVT_{t-1})}{AT_t + AT_{t-1}}$$

Lev: Leverage is the ratio of long-term debt and debt in current liabilities to the sum of long-term debt, debt in current liabilities, and stockholder's equity.

$$Lev_t = (DLTT_t + DLC_t) / SEQ_t$$

LDP: We define the dividend-price ratio as annual dividends over price.

$$LDP_t = \frac{\sum_j (RET_j - RETX_j)}{PRC_t}$$

ME: Size is the market capitalization.

Turnover: Turnover is volume over shares outstanding.

$$Turnover_t = VOL_t / SHROUT_t$$

NOA: Net operating assets are the difference between operating assets minus operating liabilities scaled by lagged total assets.

$$NOA_t = \frac{AT_t - CHE_t - IVAO_t - (AT_t - DLC_t - DLTT_t - MIB_t - PSTK_t - CEQ_t)}{AT_{t-1}}$$

NOP: Net payout ratio is common dividends plus purchase of common and preferred stock minus the sale of common and preferred stock over the market capitalization.

$$NOP_t = \frac{DVC_t + PRSTKC_t - SSTK_t}{ME_t}$$

O2P: Payout ratio is common dividends plus purchase of common and preferred stock minus the change in value of the net number of preferred stocks outstanding over the market capitalization.

$$O2P_t = \frac{DVC_t + PRSTKC_t - \Delta PSTKRV_t}{ME_t}$$

OL: Operating leverage is the sum of cost of goods sold and selling, general, and administrative expenses over total assets.

$$OL_t = (COGS_t + XSGA_t)/AT_t$$

PCM: The price-to-cost margin is the difference between net sales and costs of goods sold divided by net sales.

$$PCM_t = (SALE_t - COGS_t)/SALE_t$$

PM (WRDS Financial Ratios Firm Level database): The profit margin (operating income/sales).

Prof: We define profitability as gross profitability divided by the book value of equity.

$$Prof_t = GP_t/BE_t$$

Q: Tobin's Q is total assets plus the market value of equity minus cash and short-term investments minus deferred taxes scaled by total assets.

$$Q_t = \frac{AT_t + \frac{ME_t}{1000} - CEQ_t - TXDB_t}{AT_t}$$

Ret: Return in the month.

Ret_max: Maximum daily return in the month.

RNA (WRDS Financial Ratios Firm Level database): The return on net operating assets.

ROA (WRDS Financial Ratios Firm Level database): Return-on-assets.

ROC: ROC is the ratio of market value of equity plus long-term debt minus total assets to cash

and short-term investments.

$$ROC_t = \frac{DLTT_t + \frac{ME_t}{1000} - AT_t}{CHE_t}$$

ROE (WRDS Financial Ratios Firm Level database): Return-on-equity.

ROIC (WRDS Financial Ratios Firm Level database): Return on invested capital.

S2C: Sales-to-cash is the ratio of net sales to cash and short-term investments.

$$S2C_t = SALE_t/CHE_t$$

Sale_g: Sales growth is the percentage growth annual rate in annual sales.

$$Sale_g = SALE_t/SALE_{t-1} - 1$$

SAT: We define asset turnover as the ratio of sales to total assets.

$$SAT_t = SALE_t/AT_t$$

S2P (WRDS Financial Ratios Firm Level database): Sale-to-price is the ratio of net sales to the market capitalization.

SGA2S: SG&A to sales is the ratio of selling, general and administrative expenses to net sales.

$$SGA2S_t = XGSA_t/SALE_t$$

Spread: The bid-ask spread is the average daily bid-ask spread in the month.

Std_turnover: Standard deviation of daily turnover in the month.

Std_vol: Standard deviation of daily trading volume in the month.

Tan: Tangibility.

$$Tan_t = \frac{0.715 * RECT_t + 0.547 * INVT_t + 0.535 * PPENT_t + CHE_t}{AT_t}$$

Total_vol: Standard deviation of daily return in the month.