Developing an Actively Managed ETF Using Quantitative Strategies

Yahui Qian

Introduction

This research is undertaken to explore and design an actively managed exchange-traded fund (ETF) built on the foundation of quantitative analysis and algorithmic trading. The primary objective is to understand how data science can drive rule-based investment strategies that appeal to both institutional and retail investors. The envisioned ETF is not only a vehicle for capital appreciation but also a demonstration of how data-driven automation can democratize access to sophisticated investment tools. The knowledge base resulting from this research could be valuable to financial technology startups, personal investors interested in quantitative trading, or financial institutions seeking to enhance their algorithmic trading desks. Additionally, the insights gained will guide the development of a reproducible, algorithm-based ETF management system that can be continuously refined and deployed.

Literature Review

Much of the foundational thinking for this research draws from work in both traditional investment theory and modern quantitative finance. The pioneering framework for portfolio optimization by Harry Markowitz laid the groundwork for systematic investment strategies. His theories were further expanded by William Sharpe and subsequent scholars who examined risk-adjusted return metrics and capital asset pricing models. More recently, the success of Renaissance Technologies' Medallion Fund and firms like Two Sigma, DE Shaw, and Citadel have shown the power of algorithmic models and data-driven trading systems in achieving market-beating returns.

In the realm of ETFs specifically, Gray (2020b) and Harris (2023) have offered insights into launching and managing ETFs, including those focused on commodities. The literature on technical analysis, including the works of Meyers (2011) and Edwards et al. (2019), provides guidance on identifying signals and patterns in market data, while studies by Grinold and Kahn (2023) offer a strategic lens on active portfolio management. Risk management frameworks like those outlined by Carver (2015) and Covel (2009, 2017) are also central to defining the operational philosophy of the ETF envisioned here. Together, these sources inform a multi-faceted approach that blends financial theory, data science, and trading automation.

Methods

The research methodology centers on designing a rule-based investment strategy that can be implemented in software and tested with historical data. The first step involves identifying a universe of investable securities, potentially spanning equities, indexes, and commodities, and selecting those that align with a defined theme or market inefficiency. Technical indicators such as moving averages, momentum scores, and relative strength metrics will be computed using Python or R. These indicators serve as features in a decision system that determines entry and exit points for trades.

To simulate the ETF's performance, backtesting will be conducted using historical market data, with time series modeling playing a central role. Portfolio rebalancing, turnover rates, and transaction costs will be incorporated to ensure realistic assumptions. A trend-following strategy, incorporating both long and short positions, may be applied to manage risk dynamically. The frequency of trades will be moderate — likely monthly or biweekly — to strike a balance between capturing market trends and minimizing churn. The selected strategy

will be codified into a modular trading algorithm that can adapt to changing market conditions while adhering to transparent, auditable rules.

Results

Preliminary exploration suggests that a blend of trend-following and momentum-based strategies can provide favorable risk-adjusted returns when applied across a diversified set of equity and commodity ETFs. Backtests indicate that technical indicators such as 50-day and 200-day moving average crossovers, combined with volume spikes and volatility filters, can serve as reliable signals for trade execution. Additionally, simulations show that implementing a stop-loss and trailing stop mechanism reduces drawdowns and enhances long-term performance. The results also reveal the importance of asset class diversification, portfolios, that span multiple sectors and include non-correlated instruments (e.g., commodities and equities) tend to offer more stable returns during periods of market stress.

Conclusions

The research conducted thus far reinforces the idea that data science provides a robust foundation for building intelligent, rule-driven investment products like actively managed ETFs. It also highlights the operational complexity of implementing such a strategy in practice, especially with respect to risk control and compliance with SEC regulations. While the early findings are promising, several challenges remain, particularly regarding the handling of live data feeds, real-time execution, and the ongoing calibration of the model in response to market regime shifts.

At this stage, there are no major concerns with the overall feasibility of the term project, though attention will need to be paid to computational efficiency and interpretability of model decisions. As the project progresses, emphasis will shift toward finalizing the ETF's

investment thesis, selecting the core basket of securities, and integrating the strategy into a GitHub-hosted application. Ultimately, this research not only serves as a stepping stone toward launching a mock ETF but also deepens understanding of how data-driven decision-making can transform the landscape of personal and institutional investing.