Project Checkpoint C

The purpose of this project is to design and evaluate an actively managed ETF strategy that can be implemented algorithmically. The research is driven by the question of whether systematic, rules-based approaches can produce superior risk-adjusted returns compared with passive investing. The intended users of this work are investors who value transparency, consistency, and adaptability, as well as practitioners and academics who are interested in bridging financial theory with applied data science. Publishing the report and code in a GitHub repository allows for reproducibility and contributes to the broader quantitative finance community.

The literature that informs this work emphasizes both strategic design and methodological rigor. Classic studies such as Jegadeesh and Titman on momentum, Antonacci on dual momentum, and Asness on factor investing have shown that rules-based strategies can persistently generate excess returns. At the same time, López de Prado and Trivedi and Kyal stress the dangers of overfitting and the importance of backtesting methods that mimic real trading conditions. Together, these studies frame the dual priorities of the project: developing a practical strategy and evaluating it with robust, stress-tested methods.

The research focuses on ETF price and return data over the past ten years, from 2015 through 2024. This shorter horizon reflects a more realistic dataset for a fund concept that does not yet exist, while still capturing diverse market conditions including the COVID-19 crisis and the inflationary volatility of 2022. The methodology centers on backtesting, which is distinct from the model cross-validation techniques often used in machine learning. Cross-validation partitions a dataset into training and test folds to measure predictive accuracy, but it ignores the temporal order of financial data. Backtesting, in contrast, applies trading rules in a walk-forward manner, simulating how an investor would have acted at each point in time. This

distinction ensures that no information from the future leaks into the past and that performance is measured as if the strategy were deployed live. In this project, backtesting is performed across rolling windows to capture different phases of the market cycle, and transaction costs and management fees are incorporated to reflect investor outcomes. Monte Carlo methods are also used to generate synthetic but realistic return series, which help assess whether the strategy's apparent success is robust to different possible market paths.

Preliminary results suggest that the ETF strategy produces positive alpha relative to the S&P 500 benchmark, with moderate beta and a Sharpe ratio superior to buy-and-hold. Importantly, the backtesting process shows that the rules hold up during periods of stress. For example, volatility filters limited downside exposure during the sharp drawdowns of early 2020, while momentum signals allowed the portfolio to participate in recoveries. Including crisis periods in the backtesting record is essential, because any strategy that collapses in a downturn is unlikely to gain investor trust. By combining both historical records and simulated scenarios, the evaluation gives a more realistic picture of how the strategy may perform in live trading.