QUANTITATIVE INVESTING PROJECT

TIME SERIES ANALYSIS - STOCK PORTFOLIO STRATEGY DEVELOPMENT AND EVALUATION

Aditya Dhanotia, Jason Dormann, Shantanu Trivikram, Venkat Jishnu Channamsetti

I. Introduction

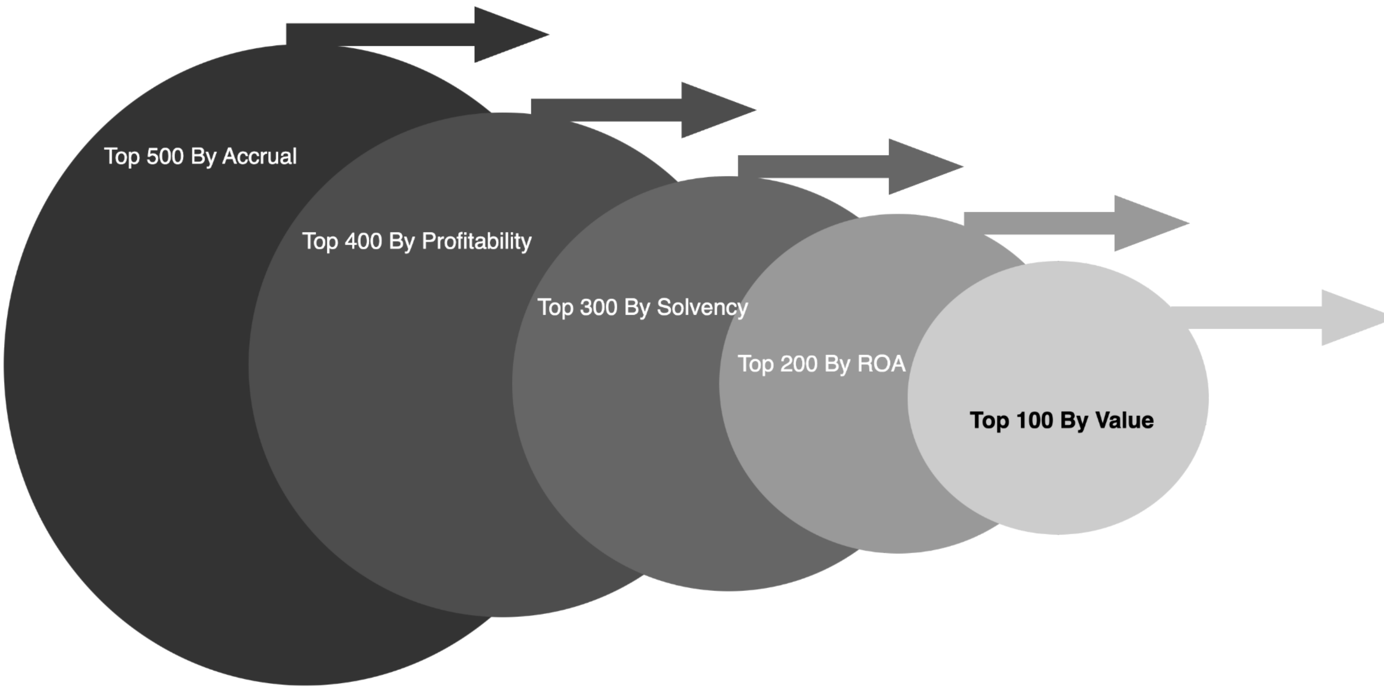
This report outlines the methodology and results of our project focused on the development and evaluation of a stock portfolio strategy. Our approach began with a comprehensive analysis of securities from the CRSP monthly stocks dataset, employing a series of filters based on financial metrics and ratios to identify undervalued stocks.

The core of our strategy involved a rigorous filtering process aimed at minimizing the risk of capital loss and identifying financially robust companies. We employed criteria such as low accruals to avoid companies potentially engaging in accounting manipulations, alongside metrics that highlighted profitability, debt servicing ability, and solvency. This systematic approach was instrumental in narrowing down our initial universe to a select group of 100 undervalued stocks.

In addition to traditional financial analysis, our strategy incorporated a novel application of machine learning techniques to estimate the expected returns of the selected stocks. This aspect of our methodology aimed to refine our portfolio weighting system, enhancing potential returns while managing the risk associated with predictive modeling. We also established a benchmark weighting system, based on an equal weighted portfolio, to evaluate the effectiveness of our machine learning-augmented approach.

The report presents detailed results of our analysis, including the performance metrics of the selected stock filters in various combinations. We also discuss the implications of these findings for the potential application and refinement of our investment strategy.

I. Filtering



We started with a universe of securities from the CRSP monthly stocks dataset. For the filtering process we employed carefully selected financial metrics/ratios to screen undervalued stocks that fulfill a certain criterion. Our approach was to first filter securities that have a risk of potential loss of capital. That meant removing securities that tend to indulge in accounting red flags. To tackle this, we filtered out stocks with high accruals.

Then, we considered stocks with high profitability, debt servicing ability, solvency, high cash flow and finally an EV multiple to select 100 undervalued stocks. Following is our reasoning to select the financial ratios for each category:

Screening 1: ACCRUALS

We wanted to be particularly cautious of companies that showcase a high net income but are low on cash flow. Such companies tend to use income recognition or expense accrual accounting practice to inflate earnings. According to our thesis of selecting from a pool of value stocks that are currently undervalued, we wanted to eliminate companies that open the door to potential financial manipulation. We empowered a ranking system where we ranked companies based on *"Accrual/Average Assets"* from low to high and picked the first 500 companies that had lowest accrual numbers aka indicating higher earnings quality, as more earnings are realized in cash.

$$\textit{Accrual} = \textit{Accruals} \,/\, \textit{Average Assets}$$

Screening 2: PROFITABILITY

The next screening step was to filter for companies who generate high cash flow. A high cash flow in itself does not mean that the company is profitable. Therefore, we looked at *"Operating Cash Flow / Total Debt"* , which signifies how much cash flow a company is generating compared to its total debt. We filtered further for the top 4=00 companies that have high OCF compared to Total Debt. High OCF is excluding any cash generated from investing or financing and tells a real picture of how strong a company's operations are. Comparing OCF with debt is a measure of debt coverage. Moreover, the 400 companies that made to the list signified a sign of financial strength, moving us closer to our goal of choosing value companies.

Cash / Debt = Operating Cash Flow / Total Debt

Screening 3: SOLVENCY

Solvency is a measure of financial risk. To further our filtering process, we wanted to ensure that the next 300 companies we choose are able to cover its interest expenses or short-term debt obligations. Here we chose a Multiple of Earnings Before Interest and Taxes to Interest and Related Expenses, *"EBIT / Interest expenses "* . EBIT focuses on the company's profitability from core operations and interest expenses are interest paid on loans and bonds. We filtered for a high ratio implying lower financial risk and a greater ability to sustain and service its debt from the 400 companies pool and selected the top 300

Interest Coverage = Earnings Before Interest and Taxes / Interest Expenses

Screening 4: RETURN ON ASSETS

Our next filter is ROA, which tells how well a company uses its assets to generate profit. A high ratio indicates that a company is utilizing its assets efficiently and generating profit. The measure of profit used here is Net Income which includes operations costs, interest, taxes and other expenses. Therefore, the measure of profitability is as purest as it can get from an income statement lens. After this step we were left with 200 companies with high profitability, debt servicing ability, efficient use of resources, and who followed non-accrual accounting practices.

ROA = Net Income / Total Assets

Screening 5: VALUE

The last step involved picking the top 100 companies ranked from a high EM to a low multiple. This step was crucial in finally picking the 100 undervalued stocks we want to create an investing strategy around. We chose an EM multiple *(EV/EBITDA)*, for our valuation metric because it better captures market mispricing compared to the traditional Book/Market value multiple. B/M doesn't capture the value premium that EV/EBITDA captures. Enterprise value captures the value of equity, debt, and cash and cash equivalents, hence, the Total Value of the company. By filtering for a low ratio we were able to carefully select undervalued companies (with high expected fundamental value), offering an opportunity for profit when the market recognizes their true value.

Valuation Multiple = Total Enterprise Value /
Earnings Before Interest, Taxes, and Depreciation

II. ML-Based Weighting

Once we have filtered our way down to 100 stocks to comprise our monthly portfolio, we need to create a strategy for the allocation of our weights. As mentioned in the introduction, we wanted to pursue a strategy that first uses machine learning to estimate, using only data that would have been publicly available at the time of portfolio construction, the expected return for the stock over the course of the next month. This, of course, is going to be a very difficult problem, as if it were not then we'd essentially be seeing statistical arbitrage as people accurately predict future returns. We feel, however, that any signal that these models are able to extract from the underlying data will allow for a slightly improved portfolio weighting system that can provide marginally improved returns. Given our skepticism about the potential accuracy of these predictive models, however, we also need to be careful that we aren't overly reliant on them and creating an irresponsible amount of variance in our weights that could greatly increase portfolio risk.

Given this framework, we first wanted to create a benchmark weighting system that would allow us to compare the performance of our portfolio weights against a more traditional scheme. For this we chose an equal weighted portfolio, with each of the 100 securities that made it through the filter receiving a weight of .01. This serves two purposes: allowing us to intrinsically compare the different filtering methods without worrying about their interaction with the modeling component and allowing us to understand the amount of additional value earned by the more complex scheme, if any.

For the machine learning portion of this two-step weighting process, we proposed two candidate models. Many more complex models could have been considered, but we will explain the considerations that were made on this issue in the next steps section of the report:

- **Market Value Regression:**

   - This strategy is as straightforward as it gets, training a linear regression model on last known market value to predict expected return for the upcoming month. This is, of course, not going to be the most accurate model that could possibly be built, but it does have a purpose to serve. This model essentially functions as an intermediate benchmark, in that we can gauge its performance relative to the equal weighting method and more complex models. If this model significantly outperforms the equal weighting method, we can feel confident that there is value to be had in the idea of an ML-based weighting scheme. Additionally, if this model is outperformed significantly by our second, more complex model, we can similarly feel confident that there is worthwhile value in investing in more accurate predictive models.

- **All-COMPUSTAT Regression:**

   - Our second strategy used every feature of the COMPUSTAT dataset, once it had been properly lagged and cleaned, in a linear regression model to predict the next month's expected return. While still not the most complex model imaginable, again more on that will come later, it still represents a significant step up in terms of model complexity.

Once expected returns have been projected for each of the 100 securities in the portfolio, we have to translate these estimates into portfolio weights before we can calculate returns and our key metrics. We knew that we wanted to constrain our range of possible weights such that some extreme predicted value did not lead to an imbalance portfolio. We settled on two variations of the same core strategy:

- Rank-Proportional Weighting
   - In this strategy, a security is first given a rank inverse to its expected rate of return over the next month, such that a rank of 0 is given to the security with the lowest expected return, and a rank of 99 is given to the security

with the highest expected return. We then give each security a weight that is equal to its own rank divided by the sum of all ranks. This leads us to have weights uniformly distributed between ~2% and 0%. We settled on this strategy as it provided lower variance than other candidates, while still providing the potential for overweighting stronger investments.

- Double-Leveraged Rank-Proportional Weighting
  - This strategy functions the same as the previous, except at the final step, we take double leverage and multiply each weight by 2, leading us to have a total of 200% invested. This, of course, is a strategy that carries a much more significant amount of risk, but it also provides a greater emphasis on the predicted returns as now the gap in weight between the highest and lowest weighted securities is ~4% instead of ~2%. This means that we will observe a slightly different Sharpe ratio which can help us make inferences about whether or not future strategies should provide larger gaps in the weights based on expected return predictions.

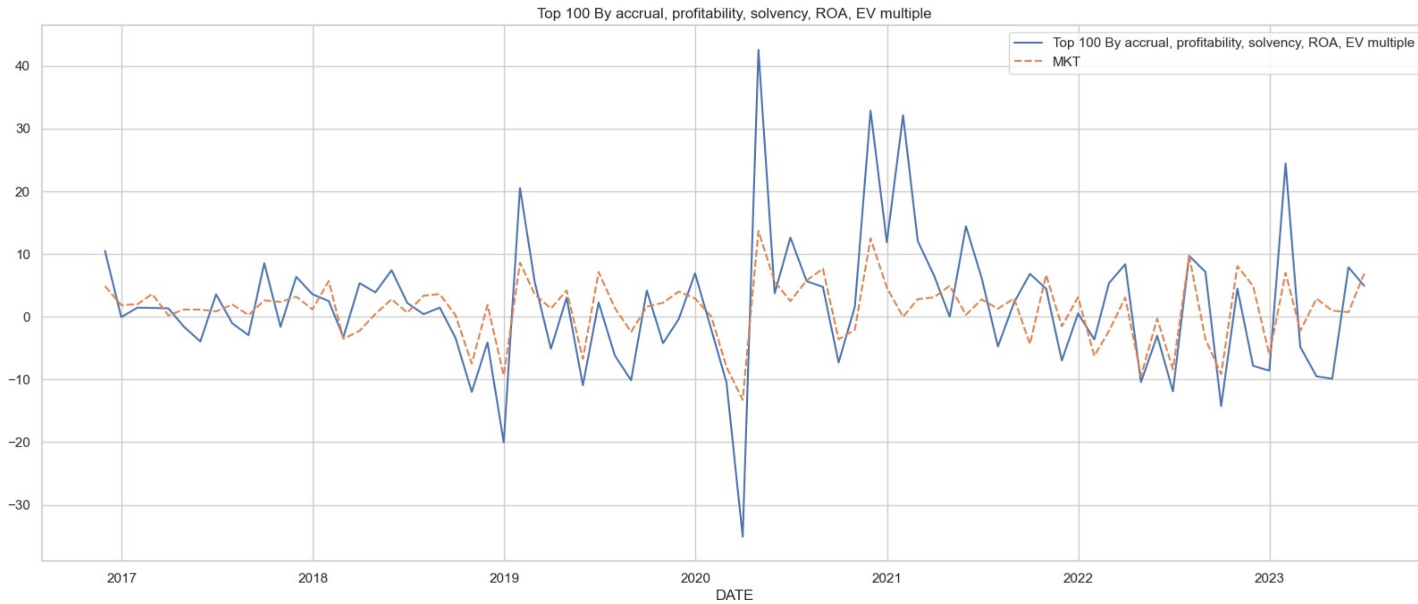
III. Results

In the results analysis, we employed various financial metrics and ratios to systematically filter companies, utilizing market returns as the benchmark. Across different combinations of filters, we calculated key performance metrics such as Excess Returns, Sharpe Ratio, Max Drawdown, Downside Deviation, Sortino Ratio, Beta, and Alpha. The outcomes of our analysis are detailed below for select filter combinations:

1. Top 100 companies filtered by Accrual, Profitability, Solvency, Return on Assets, Value:

The Market Value Regression with Double Leverage Rank stood out in this group, demonstrating an Alpha of 0.32. This indicates that, after adjusting for risk, the portfolio achieved a return 0.32% higher than expected, showcasing the effectiveness of the selected financial filters. The substantial 48% excess return suggests a robust performance, and the 65% volatility indicates the level of risk associated with this investment strategy. The combination of these factors underscores the portfolio's potential for generating attractive risk-adjusted returns.

Filter	Model	Weighting	Annualized Excess Return Over Market	Annualized Volatility of Excess Return	Sharpe Ratio	T-Stat Excess Return	Max Drawdown	Downside Deviation	Sortino Ratio	CAPM Beta	CAPM Alpha
Top 100 By accrual, profitability, solvency, ROA, EV multiple		Equal	10.8001%	27.0385%	0.4749	1.9363	65.3589%	0.0609	0.2367	1.6608	0.0651
	Market Value Regression	Proportional Rank	20.4403%	28.1985%	0.7121	3.5139	56.9524%	0.0566	0.3966	1.7401	0.1563
		Double Leverage Rank	48.924%	64.7423%	0.7329	3.6633	87.6699%	0.1123	0.4112	3.4758	0.3285
	All COMPUSTAT Regression	Proportional Rank	11.2585%	26.8035%	0.4972	2.0362	66.6101%	0.0596	0.2482	1.5856	0.0746
		Double Leverage Rank	30.5604%	61.0234%	0.5193	2.4277	92.6723%	0.1182	0.2612	3.1668	0.1649



2. Top 100 companies by Accrual, Solvency, Return on Assets, Profitability, and Value:

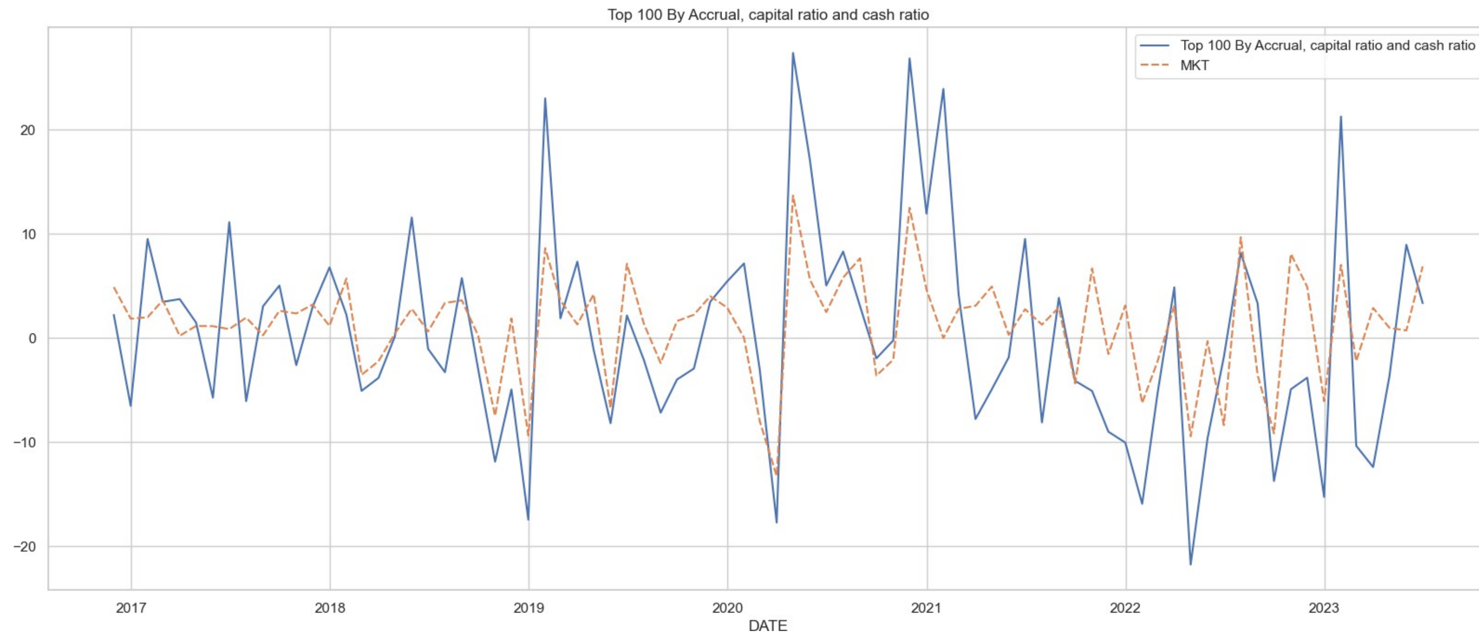
In this scenario, the Market Value Regression with Double Leverage Rank delivered an Alpha of 0.28, signaling a positive risk-adjusted performance. The 44% excess return and 61% volatility provide a balanced perspective on both returns and associated risks. This result indicates that the selected financial filters effectively identified companies that outperformed market expectations.

Filter	Model	Weighting	Annualized Excess Return Over Market	Annualized Volatility of Excess Return	Sharpe Ratio	T-Stat Excess Return	Max Drawdown	Downside Deviation	Sortino Ratio	CAPM Beta	CAPM Alpha
Top 100 By accraul, solvency, roa, profitability, EV multiple	Market Value Regression	Equal	9.108%	26.5516%	0.4330	1.6629	65.6356%	0.0605	0.2149	1.6581	0.0483
		Proportional Rank	18.2095%	26.6597%	0.6758	3.3111	54.7119%	0.0564	0.3650	1.7204	0.1353
		Double Leverage Rank	44.4625%	61.9245%	0.6973	3.4807	86.3875%	0.1121	0.3788	3.4363	0.2864
	All COMPUSTAT Regression	Proportional Rank	9.7339%	25.8292%	0.4669	1.8269	68.9748%	0.0585	0.2312	1.5557	0.0612
		Double Leverage Rank	27.5113%	59.0791%	0.4895	2.2574	93.1019%	0.1161	0.2441	3.1070	0.1383

3. Top 100 companies filtered by Accrual, Capital Ratio, and Cash Ratio:

The Market Value Regression with Double Leverage Rank in this group exhibited an Alpha of 0.30, indicating a positive excess return after accounting for risk. The substantial 44% excess return coupled with a volatility of 69% implies a potentially lucrative but riskier investment. The Sharpe ratio of 0.64 further reinforces the risk-return profile, with a higher Sharpe ratio suggesting a better risk-adjusted performance compared to alternative investments with similar volatility.

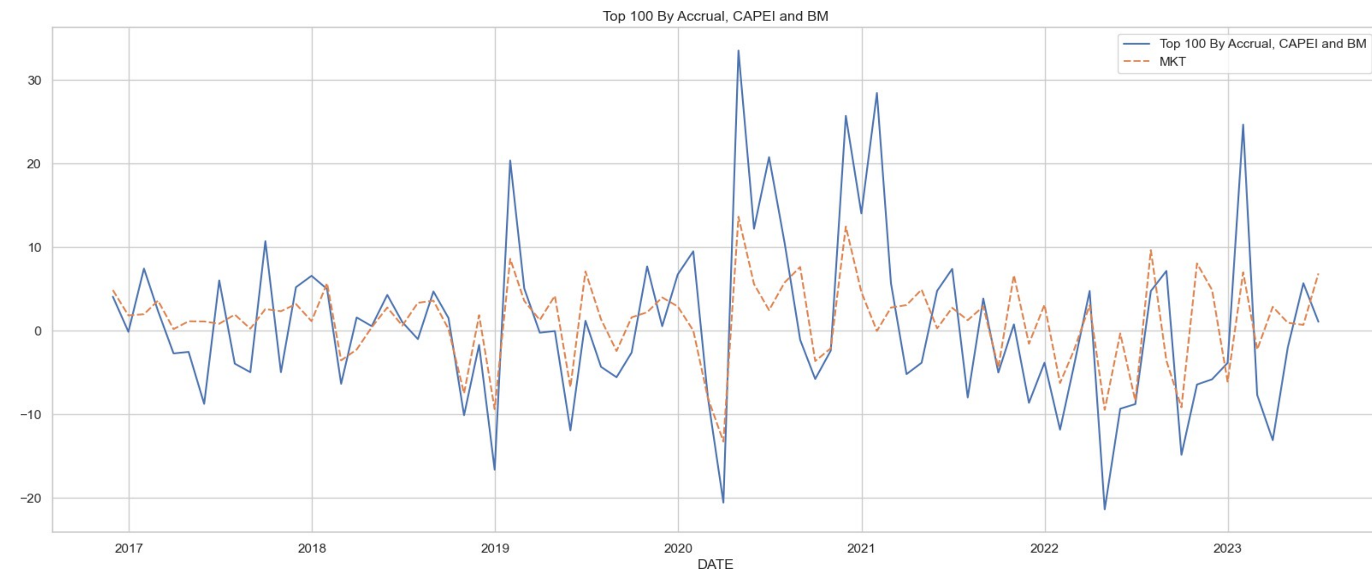
Filter	Model	Weighting	Annualized Excess Return Over Market	Annualized Volatility of Excess Return	Sharpe Ratio	T-Stat Excess Return	Max Drawdown	Downside Deviation	Sortino Ratio	CAPM Beta	CAPM Alpha
Top 100 By Accrual, capital ratio and cash ratio		Equal	0.7062%	31.1607%	0.1840	0.1099	79.549%	0.0685	0.0876	1.5920	-0.0314
	Market Value Regression	Proportional Rank	18.1321%	31.1752%	0.6256	2.8195	71.1099%	0.0619	0.3315	1.6272	0.1406
		Double Leverage Rank	44.3077%	69.1545%	0.6456	3.1059	97.4677%	0.1229	0.3445	3.2501	0.2969
	All COMPUSTAT Regression	Proportional Rank	1.3455%	30.3662%	0.2061	0.2148	74.5842%	0.0652	0.1002	1.5246	-0.0206
		Double Leverage Rank	10.7345%	66.9143%	0.2266	0.7777	98.5388%	0.1293	0.1110	3.0448	-0.0255



4. Top 100 companies filtered by Accrual, CapEI, and BM:

In this distinctive set of filters, the Market Value Regression with Double Leverage Rank displayed a notably high Alpha of 0.47. This implies a superior risk-adjusted return, reinforcing the effectiveness of the chosen financial criteria. The impressive 60% excess return and 67% volatility highlight the potential for substantial returns, albeit with a higher level of associated risk. The Sharpe ratio of 0.90 signifies a compelling risk-adjusted performance, making this combination particularly attractive from an investment perspective.

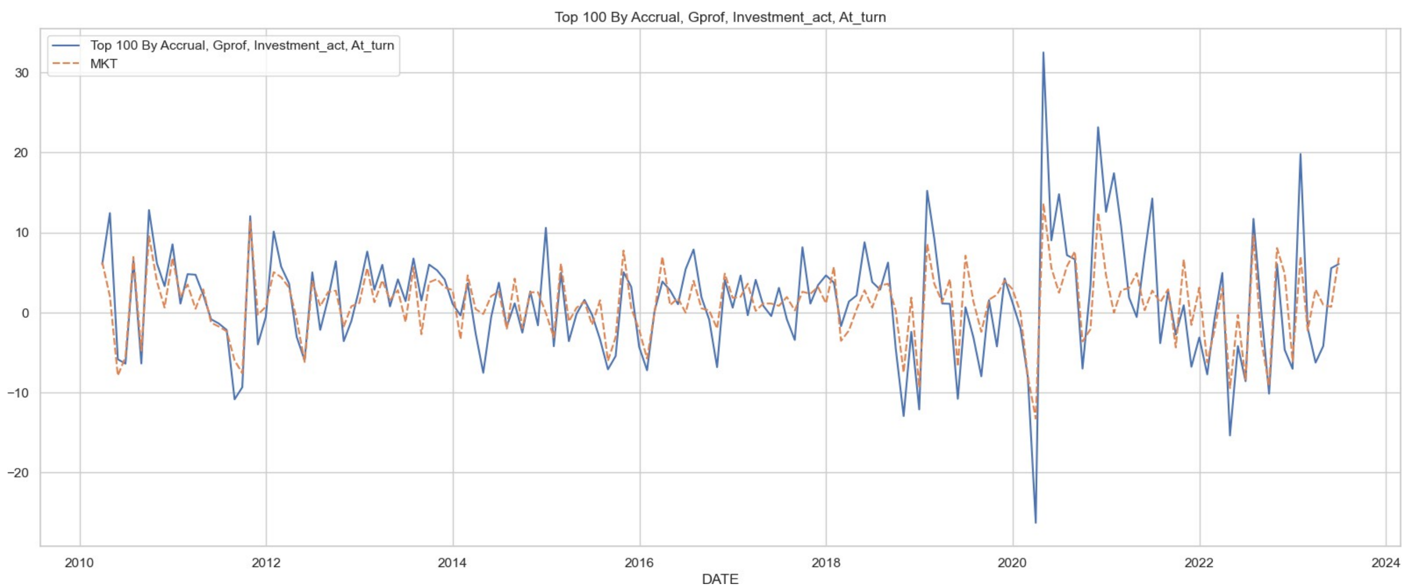
Filter	Model	Weighting	Annualized Excess Return Over Market	Annualized Volatility of Excess Return	Sharpe Ratio	T-Stat Excess Return	Max Drawdown	Downside Deviation	Sortino Ratio	CAPM Beta	CAPM Alpha
Top 100 By Accrual, CAPEI and BM		Equal	12.0376%	23.8697%	0.5697	2.4447	61.2402%	0.0528	0.2925	1.4531	0.0909
	Market Value Regression	Proportional Rank	23.2776%	21.5873%	0.9553	5.2273	51.4762%	0.0469	0.5290	1.4857	0.2012
		Double Leverage Rank	54.5986%	51.1288%	0.9809	5.1767	81.2049%	0.0929	0.5480	2.9670	0.4182
	All COMPUSTAT Regression	Proportional Rank	11.2748%	23.3764%	0.5512	2.3381	61.568%	0.0529	0.2799	1.4613	0.0828
		Double Leverage Rank	30.5929%	53.9344%	0.5758	2.7497	92.6007%	0.1048	0.2949	2.9183	0.1813

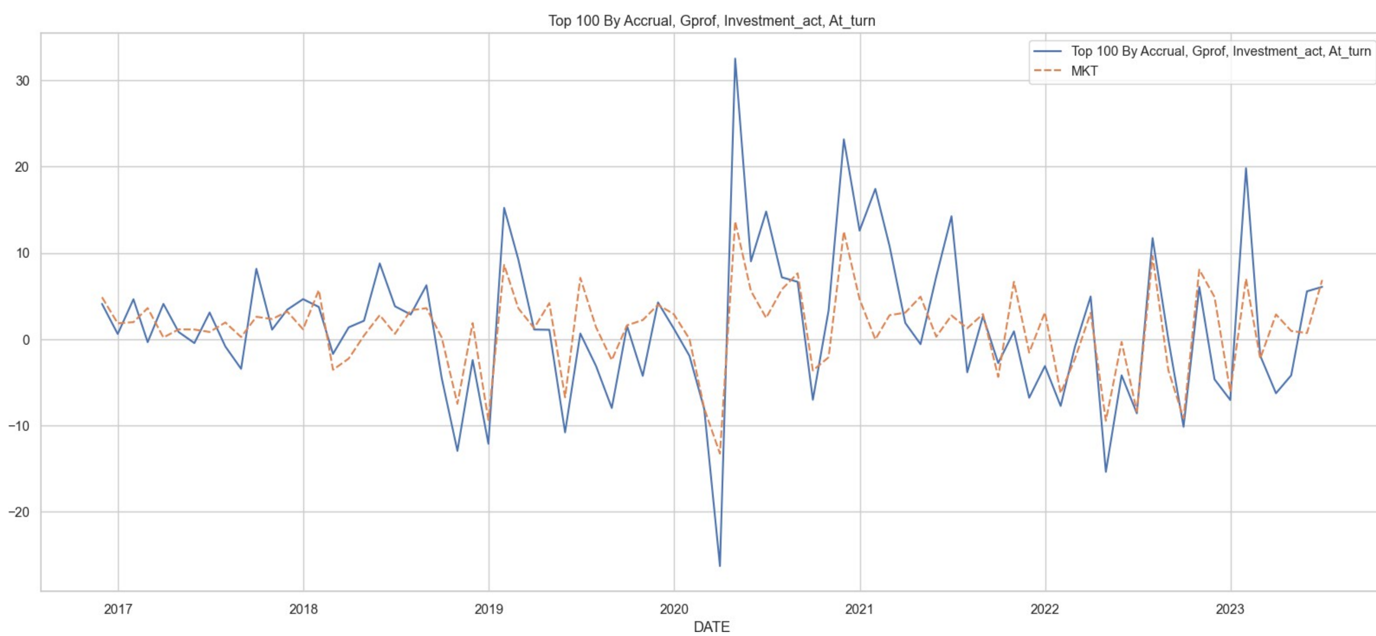


5. Top 100 companies filtered by Accrual, Profitability, Investments, Value:

Within this filter combination, the Market Value Regression with Double Leverage Rank showcased an Alpha of 0.43, indicating a strong risk-adjusted performance. The 58% excess return and 67% volatility suggest a potentially lucrative but riskier investment strategy. The favorable Sharpe ratio supports the notion of a well-balanced risk-return profile, making this combination appealing for investors seeking both substantial returns and effective risk management.

Filter	Model	Weighting	Annualized Excess Return Over Market	Annualized Volatility of Excess Return	Sharpe Ratio	T-Stat Excess Return	Max Drawdown	Downside Deviation	Sortino Ratio	CAPM Beta	CAPM Alpha
Top 100 By Accrual, Gprof, Investment_act, At_turn		Equal	3.6234%	30.7166%	0.2636	0.5718	80.4628%	0.0634	0.1330	1.5326	0.0016
	Market Value Regression	Proportional Rank	26.3355%	29.0454%	0.8864	4.3954	53.7996%	0.0496	0.5516	1.5300	0.2289
		Double Leverage Rank	60.7144%	64.5984%	0.9076	4.5562	83.8521%	0.0984	0.5692	3.0557	0.4736
	All COMPUSTAT Regression	Proportional Rank	3.8246%	28.0888%	0.2881	0.6601	76.855%	0.0586	0.1467	1.4635	0.0081
		Double Leverage Rank	15.6926%	62.3124%	0.3099	1.2208	97.5048%	0.1163	0.1590	2.9226	0.0321





These detailed analyses underscore the nuanced outcomes of different filter combinations, providing investors with insights into the risk-return profiles of various investment strategies. Each result emphasizes the importance of carefully selecting and combining financial filters to identify companies with the potential for mispricing and attractive investment opportunities.

IV. Potential Next Steps

There are two pieces which were outside the scope of this project that we feel could lead to even better portfolio performance for this strategy, as well as other potential concerns that would need to be researched before any sort of implementation of the trading procedure in practice.

- Rigorous Testing of Complex Machine Learning Methods: Our focus in this project was more surrounding how to translate the predictions of expected return into portfolio weights, so we did not dig deeply into more complex models, feature engineering, and so on. It is obviously true, however, that there's a marginal increase in the performance of this multi-pronged strategy to be made by developing more accurate predictions. We ultimately decided, however, that this was a tangential addition to the strategy we were laying out, and that we also lacked the computing power necessary to train and validate a wide variety of complex neural network models each month, though that would likely lead to optimal portfolio performance on the highly nonlinear and noisy patterns present in financial data.

- Analysis of Transaction Costs and Portfolio Rebalancing: Another avenue that would be worth further exploring before implementation of this trading strategy is a measure of how frequently certain securities come into or out of the portfolio based on each candidate filter, and map these to an effect on transaction costs. It may be the case that a more complex filtering system that has some sort of rule that enforces a "memory" property similar to an LSTM network could help optimize returns while keeping transaction costs and rebalancing complexity low.

V. References

Fitzgibbons, S., Friedman, J. A., Pomorski, L., & Serban, L. (2016). Long-Only Style Investing: Don't Just Mix, Integrate. AQR Capital Management. Retrieved from https://www.aqr.com/Insights/Research/White-Papers/Long-Only-Style-Investing.

Gray, W., Vogel, J., Foulke, S., & Ulrich, E. (2022). Long-Only Value Investing: A Historical Perspective. Alpha Architect. Retrieved from https://alphaarchitect.com/wp-content/uploads/2022/11/AA-JBISFactorInvesting22LongOnlyValueInvesting.pdf.