MGTF 430: Behavioral Finance – Trading Strategy Research

Justin Yu Chi Hsu, Patricio Harris, Tushar Mohan, Xiaoyu Zhao

Harnessing Behavioral Biases with Cutting-Edge Quantitative Trading Strategies:

Integrating Market Psychology & Neural Network

Introduction

Today's dynamic financial markets demand complex quantitative trading strategies as essential tools for investors seeking to capitalize on market inefficiencies. The primary objective of our strategy is to exploit various behavioral biases to achieve consistent returns. Our research process involved a thorough examination of various factors that influence market behavior. After identifying and researching these individual factors, we transitioned to combining them through sophisticated portfolio optimization techniques. This phase used mean-variance optimization (MVO) and active reallocation based on the anticipated performance of each factor. Key factors include: *(See appendix for detailed composition and results)*

- *Familiarity Bias(FOMO or Fear of Missing Out):* This factor captures the impact of the industry the firm operates in and momentum-driven trading, where investors buy into trends to avoid missing potential gains.
- *Valuation Bias*: This factor exploits mispricings by focusing on the convergence of valuation metrics, targeting stocks expected to revert to their mean earnings-price ratio.
- *%Sales - %Inventory*: This is a yearly factor since we only get yearly data of sales and inventory from CompuStat. percentage of Sales/Inventory means a percentage change in the variable from its average over the past two years.
- *Technical Analysis:* Utilizing neural networks, we generate predictive rankings from various technical indicators, allowing us to identify stocks with high potential returns.

By combining these factors, our strategy seeks to capitalize on a number of influences driving market movements. Furthermore, the optimization techniques ensure that our portfolio remains responsive to changing conditions.

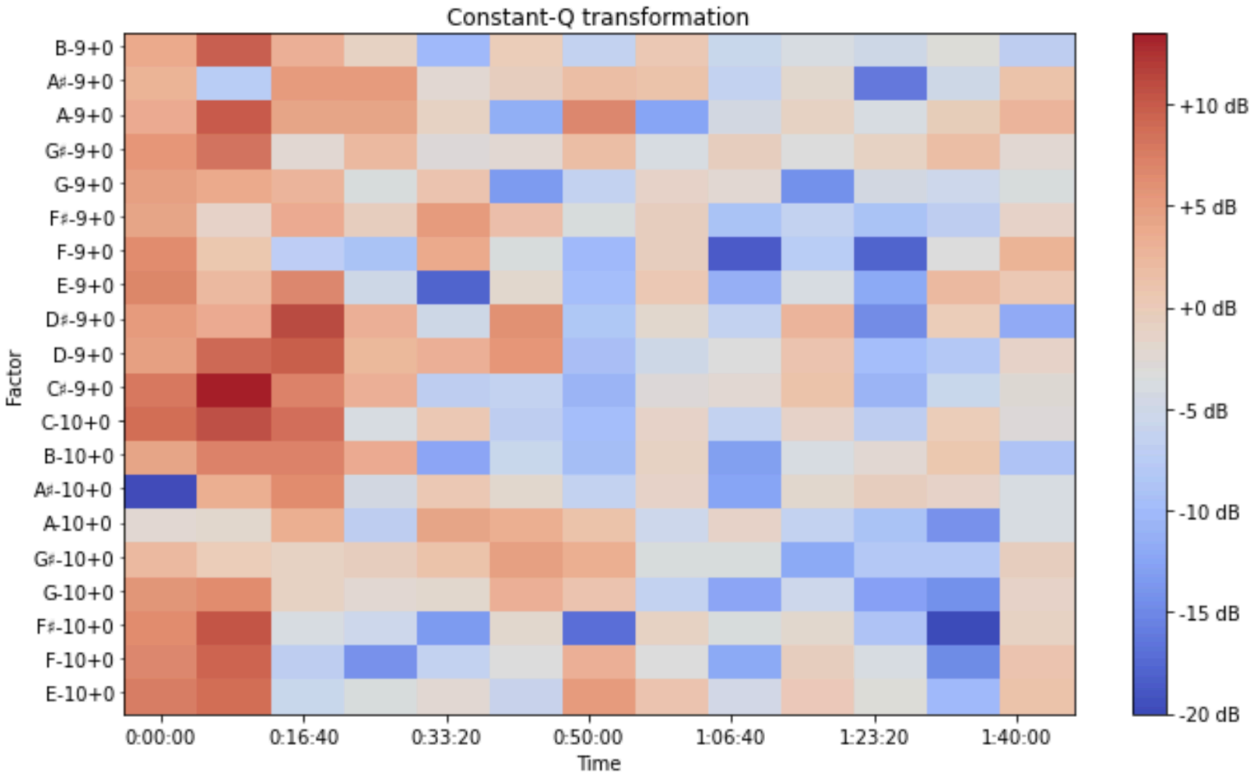
Data description

Factor Performance: *Constant Q transformation*

Constant Q Transform (CQT) is a method used to transform a signal from the time domain to the frequency domain. Unlike the Fourier Transform, it offers a more adaptive frequency resolution. CQT is particularly suitable for analyzing nonlinear frequencies in audio signals, such as musical signals, as it provides frequency resolution on a logarithmic scale, making it more effective for analyzing signals with different frequency components.

Although CQT is mainly used in audio signal processing, its principles can also be applied to other fields, such as financial time series analysis. When analyzing stock returns, CQT can be used to capture periodic or frequency characteristics at different time scales

.



Application Steps:

1. **Parameter Selection**:We select 20 bins, which means we will extract 20 different dimensional attributes of the time series. We set the sampling rate and hop length to 1, ensuring that we capture all available information. The remaining parameters are set to their default values.
2. **Transform Calculation**: Import our stock return series into a CQT package LIBROSA to transform the data from the time domain to the frequency domain, capturing its factors components.
3. **Create Rolling Window**: After getting the CQT data at each time point, we transform our data into rolling windows so we can use the window to predict future stock return.
4. **Standardize Data Format:** We set our output data as a standard 4D array. Each dimension represents time, rolling window, different CQT dimensions, and different stocks, respectively, to facilitate subsequent import into the neural network.

Individual Factor Performance: *Familiarity Bias / FOMO*

The first factor we examined aims to capitalize on Familiarity bias and the fear of missing out that is prevalent in markets influenced by the industry the company operates. Using the SICS classification and money trades (price times volume) for each stock, we rank industries into deciles based on unusual money trades. This strategy captures gains based on the familiarity and fear of missing out on investors in certain industries.

Methodology:

- Unusual Money Trades: This factor uses 12-month rolling z-scores of "price times volume" to identify industries showing signs of FOMO or familiarity bias.
- Ranking and Classification: Stocks are cross-sectionally ranked into deciles based on their SIC industry classification's z-score. Higher deciles represent stocks with greater FOMO or familiarity bias activity.
- Long-Short Strategy: The strategy involves going long on stocks in the top decile (Q10) and shorting stocks in the bottom decile (Q1), capitalizing on positive feedback trading due to FOMO or familiarity bias.

Results:

- Mean Returns: The mean returns increase with higher deciles, indicating that stocks with more significant FOMO/familiarity activity tend to perform better. The monthly mean return for the top decile (Q10) is approximately 1.4658%, while the mean return for the bottom decile (Q1) is about 0.9228%.
- Sharpe Ratio: The Sharpe ratio of 0.649725 suggests a moderate return-to-risk ratio. While not extremely high, it still indicates a positive performance relative to the risk taken.
- Holding Period Return (HPR): The HPR chart shows the returns for holding the stocks for 3 months and 1 year. The strategy captures significant market events, such as the Dot Com boom, demonstrating its ability to identify profitable trends.
- Decile Performance: The chart showing the performance of each decile portfolio highlights the increasing returns with higher deciles. The resetting of the scale to $1 at each recession period makes the divergence between deciles more apparent and underscores the strategy's resilience.

Performance For Each Decile Portfolios

Q	count	mean	std	min	25%	50%	75%
1	635.0	0.009228	0.058820	-0.271175	-0.023200	0.011464	0.041640
2	635.0	0.009598	0.057447	-0.294566	-0.020535	0.009982	0.041750
3	635.0	0.009690	0.056468	-0.282916	-0.022443	0.013348	0.039710
4	635.0	0.011032	0.056700	-0.274451	-0.020165	0.014366	0.042973
5	635.0	0.011326	0.055537	-0.268447	-0.017836	0.012668	0.045454
6	635.0	0.012509	0.055616	-0.285644	-0.018442	0.016646	0.045185
7	635.0	0.011780	0.053400	-0.285642	-0.016215	0.014752	0.043348
8	635.0	0.012982	0.053340	-0.267413	-0.017396	0.014205	0.044581
9	635.0	0.014157	0.053824	-0.300751	-0.016531	0.015993	0.047044
10	635.0	0.014658	0.053313	-0.284509	-0.014479	0.014350	0.046949

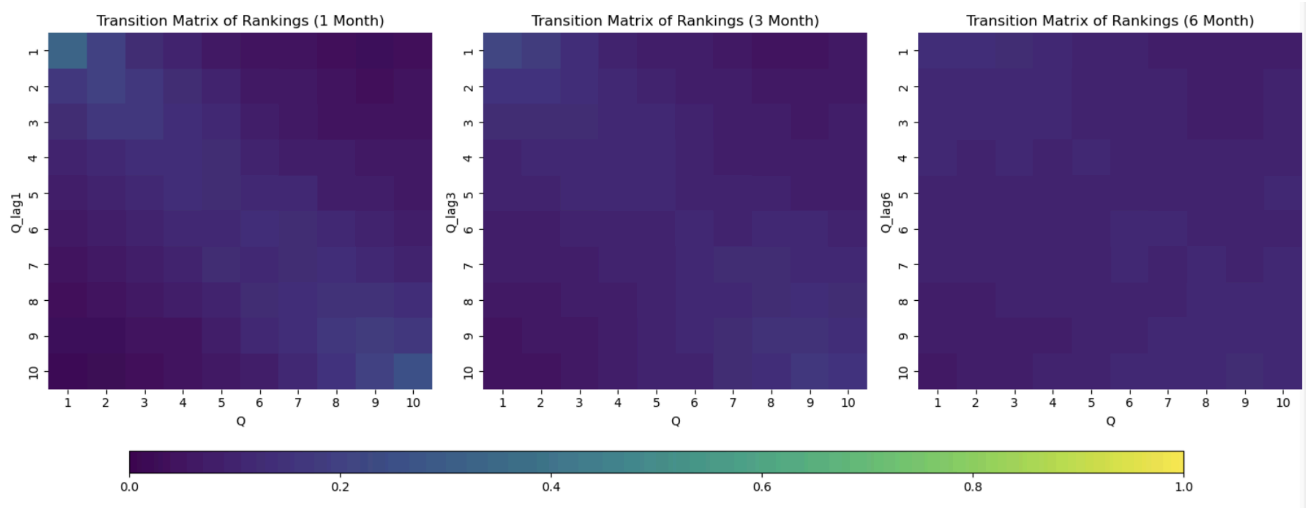
Q	max
1	0.261263
2	0.304429
3	0.268111
4	0.289331
5	0.238012
6	0.222234
7	0.257152
8	0.248503
9	0.188625
10	0.269451

```
---------- Performance of Factor Portfolio----------
count     635.000000
mean        0.005431
std         0.028953
min        -0.127621
25%        -0.008667
50%         0.004605
75%         0.019317
max         0.292025
tstat       4.726350
sharpe      0.649725
Name: RET, dtype: float64
```

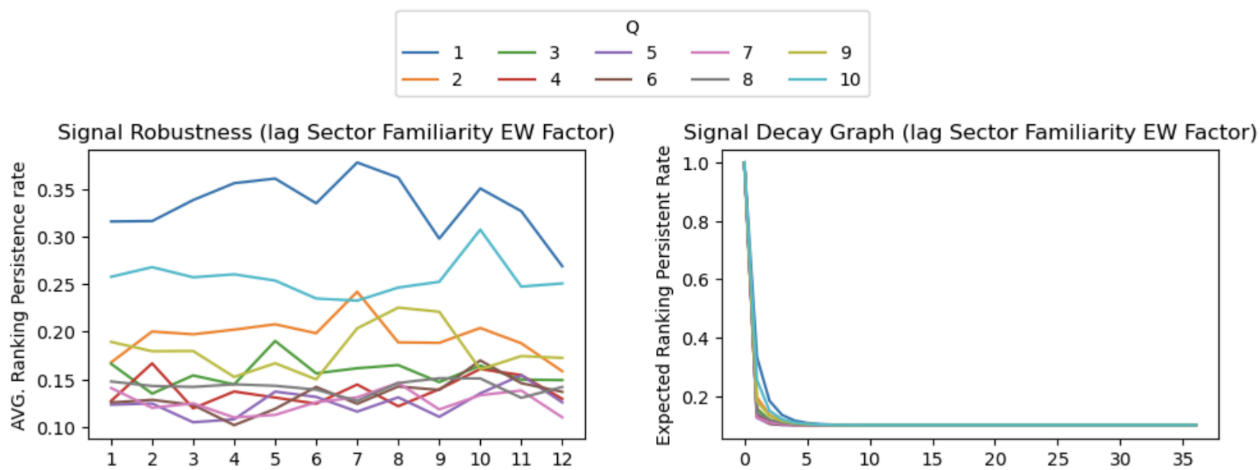
Alpha Analysis by Recession Separation

	count	mean	std	min	25%	50%	75%	max	tstat	sharpe ratio	MaxDD	α	p_value (α)	β_Market	p_value (β_Market)	R-squared
1971-01 ~ 1973-10	34	0.0111	0.0219	-0.0394	-0.0029	0.0123	0.0243	0.0630	2.9684	1.7635	-0.0396	0.0120	0.0010	-0.3097	0.0033	0.2390
1973-11 ~ 1975-03	17	0.0006	0.0340	-0.0547	-0.0199	-0.0023	0.0197	0.0584	0.0737	0.0620	-0.0758	-0.0036	0.6175	-0.2541	0.0137	0.3417
1975-04 ~ 1979-12	57	0.0013	0.0187	-0.0426	-0.0108	0.0030	0.0125	0.0450	0.5168	0.2371	-0.1194	0.0009	0.7240	0.0656	0.2796	0.0212
1980-01 ~ 1980-07	7	-0.0012	0.0354	-0.0324	-0.0206	-0.0193	0.0067	0.0709	-0.0923	-0.1209	-0.0776	-0.0030	0.8445	0.1235	0.6045	0.0575
1980-08 ~ 1981-06	11	0.0194	0.0235	-0.0398	0.0147	0.0226	0.0302	0.0553	2.7275	2.8488	-0.0398	0.0187	0.0279	0.1735	0.3687	0.0905
1981-07 ~ 1982-11	17	0.0275	0.0309	-0.0248	0.0100	0.0248	0.0620	0.0735	3.6674	3.0813	-0.0230	0.0272	0.0019	-0.1879	0.1692	0.1221
1982-12 ~ 1990-06	91	0.0081	0.0192	-0.0407	-0.0028	0.0099	0.0200	0.0658	4.0283	1.4628	-0.0732	0.0080	0.0002	0.0067	0.8777	0.0003
1990-07 ~ 1991-03	9	-0.0030	0.0228	-0.0311	-0.0172	-0.0061	0.0065	0.0461	-0.4008	-0.4628	-0.0608	-0.0025	0.7458	-0.1783	0.2356	0.1939
1991-04 ~ 2001-02	119	0.0121	0.0404	-0.1276	-0.0039	0.0099	0.0260	0.2920	3.2754	1.0401	-0.1555	0.0115	0.0030	0.0814	0.3763	0.0067
2001-03 ~ 2001-11	9	-0.0004	0.0576	-0.0966	-0.0344	-0.0013	0.0371	0.0900	-0.0187	-0.0216	-0.1341	-0.0066	0.6461	-0.6882	0.0201	0.5618
2001-12 ~ 2007-11	72	0.0041	0.0245	-0.0895	-0.0083	0.0024	0.0184	0.0695	1.4041	0.5732	-0.2174	0.0056	0.0398	-0.2956	0.0002	0.1773
2007-12 ~ 2009-06	19	-0.0098	0.0541	-0.1162	-0.0337	-0.0147	0.0166	0.1110	-0.7919	-0.6294	-0.2734	-0.0188	0.0909	-0.4448	0.0056	0.3719
2009-07 ~ 2020-01	127	0.0002	0.0183	-0.0627	-0.0085	0.0004	0.0072	0.0559	0.0974	0.0299	-0.1232	0.0018	0.2840	-0.1456	0.0008	0.0860
2020-02 ~ 2020-06	5	0.0213	0.0404	-0.0363	0.0146	0.0189	0.0328	0.0764	1.1782	1.8253	-0.0363	0.0204	0.0591	-0.3525	0.0155	0.8923
2020-07 ~ 2023-11	41	-0.0016	0.0207	-0.0474	-0.0114	-0.0016	0.0123	0.0473	-0.5057	-0.2736	-0.1948	-0.0009	0.7725	-0.0748	0.2362	0.0358

Signal Analysis



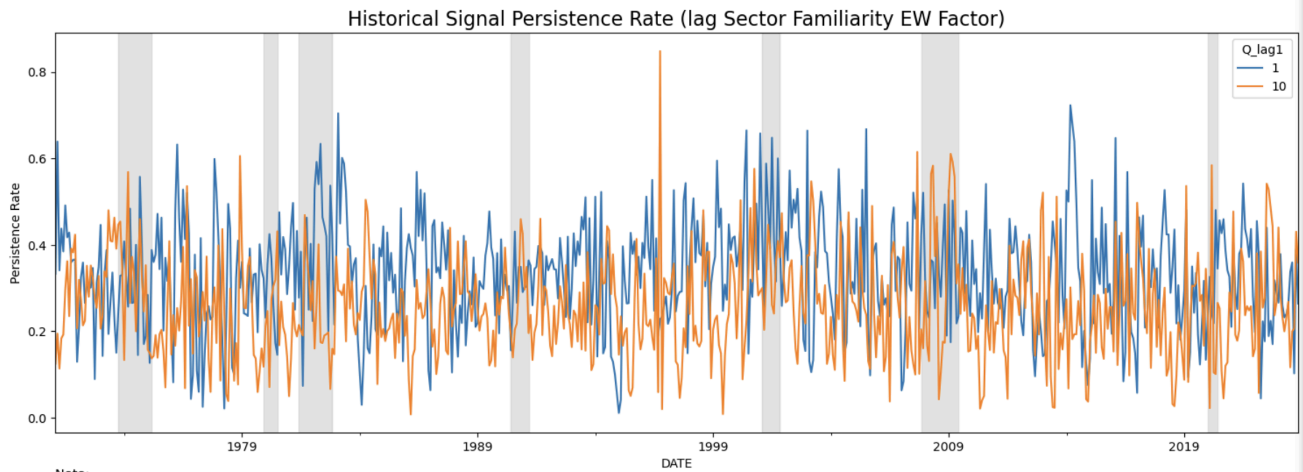
The above 3 graphs depict how inconsistent the ranks of the stocks are for 1 month, 3 months and 6 months. If the stocks' ranks were persistent, we would have seen a lighter shade along the diagonal but it is pretty dark which shows negligible persistence.



Note:
The "Ranking Persistence Rate" measures the percentage of rankings that remain the same compared to the previous ranking. In other words, it shows how often a ranking position is unchanged from its previous position.

Note:
The Expected Signal Persistent rate is calculate through transition matrix(1month).

The first graph shows the percentage of rankings that remain consistent over the calendar months on an average for each decile stock. More specifically, it depicts how consistent are the ranks of the stocks across the months on average across the years.
The graph on the right depicts the probability of a stock holding its rank over the months, and it is clear that apart from the top 2 decile stocks, the drop in probability is drastically high. The persistence probability drops to 0.3 in the first 2 months itself.



Note:
The "Ranking Persistence Rate" measures the percentage of rankings that remain the same compared to the previous ranking.
In other words, it shows how often a ranking position is unchanged from its previous position.

Individual Factor Performance: *Valuation Bias*

Following the first factor, we incorporated Valuation Bias to identify and exploit opportunities where stocks are mispriced relative to their intrinsic values. This factor targets securities with drastic changes in valuation, likely due to overreaction or underreaction, and aims to capitalize on their expected mean reversion based on historical valuation metrics.

By focusing on the convergence of valuation metrics, the strategy aims to capture gains from the natural correction of these mispricings.

Methodology:

- Valuation Metrics: We use rolling mean of fundamental valuation metrics earnings-to-price (E/P) ratio to assess the intrinsic value of stocks. We calculate the moving z-score of the E/P ratio for each stock and then calculate the first difference of the z-score.
- Ranking and Classification: Stocks are ranked by the first difference of the rolling z-score. The top decile (Q10) indicates the most likely undervalued stocks, while the bottom decile (Q1) includes the most likely overvalued stocks.
- Trading Signals: The strategy involves going long on the top deciles and shorting the bottom deciles, anticipating a mean reversion in their valuations.

Results:

- Mean Returns: The Q10 portfolio has a monthly mean return of 1.863%, while the Q1 portfolio has a mean return of 0.5478%, indicating effective identification of opportunities.
- Sharpe Ratio: A strong Sharpe ratio of 1.697844 highlights a favorable return-to-risk ratio, reinforcing the strategy's robustness.
- Portfolio Growth: The "Performance of each portfolio ($1 invested)" chart shows significant growth in the higher quantiles, particularly in the top decile (Q10), which demonstrates substantial gains over time.
- Consistency: The holding period return chart reflects consistent performance across different market conditions, with fewer extreme spikes compared to other factors.

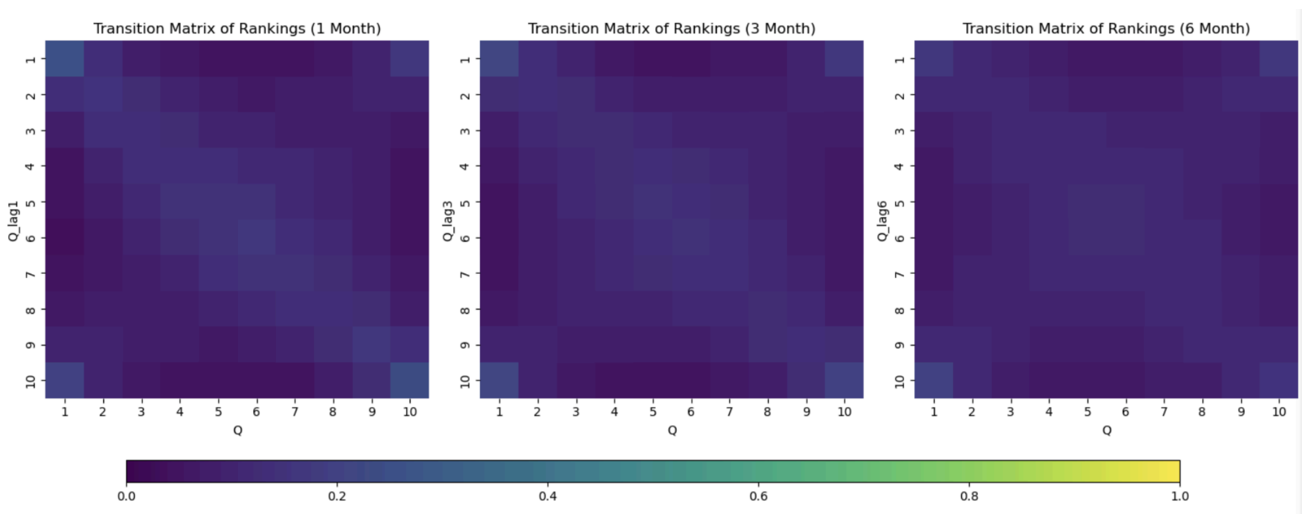
Performance For Each Decile Portfolios

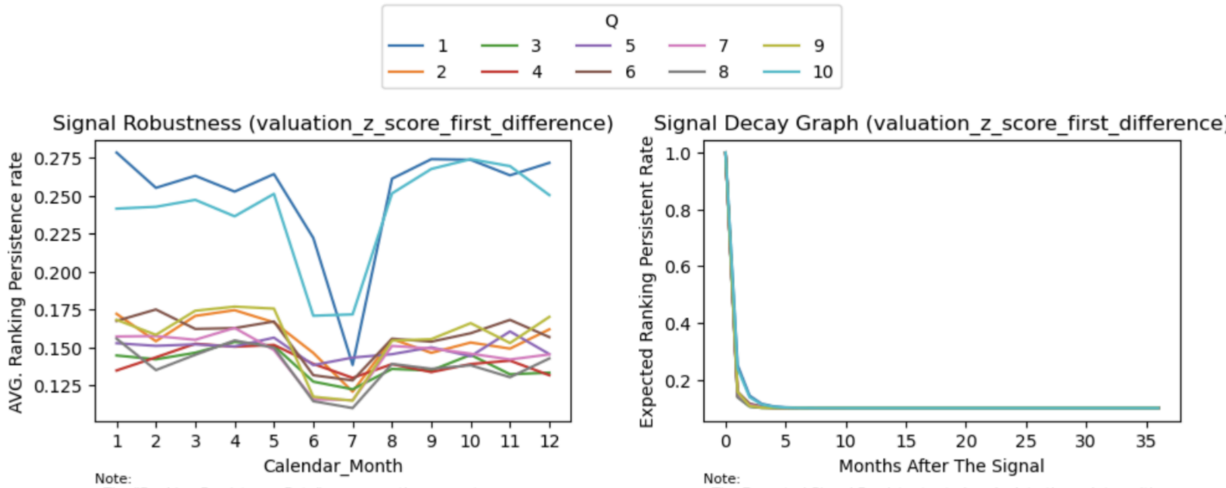
```
---------- Performance of Each Quantiles Portfolios----------
    count      mean       std       min       25%       50%       75%  \
Q
1   599.0  0.005478  0.049818 -0.284444 -0.021544  0.007224  0.037311
2   599.0  0.008761  0.049939 -0.281879 -0.017253  0.012458  0.038450
3   599.0  0.010915  0.051770 -0.281797 -0.017125  0.013969  0.041753
4   599.0  0.012211  0.052231 -0.268976 -0.015450  0.016597  0.042729
5   599.0  0.013264  0.054385 -0.255234 -0.017984  0.016912  0.043345
6   599.0  0.014699  0.054508 -0.259767 -0.014892  0.016680  0.046064
7   599.0  0.015357  0.055924 -0.252568 -0.015770  0.018878  0.046467
8   599.0  0.016897  0.054830 -0.260193 -0.013541  0.018774  0.050935
9   599.0  0.017498  0.057984 -0.267636 -0.014270  0.018434  0.047868
10  599.0  0.018631  0.057990 -0.261268 -0.015466  0.019114  0.050903

         max
Q
1   0.225438
2   0.209684
3   0.208115
4   0.248980
5   0.321677
6   0.355622
7   0.381272
8   0.253920
9   0.397448
10  0.429400
---------- Performance of Factor Portfolio----------
count     599.000000
mean        0.013153
std         0.026835
min        -0.089610
25%        -0.001039
50%         0.010889
75%         0.023373
max         0.245542
tstat      11.995560
sharpe      1.697844
Name: RET, dtype: float64
```

Alpha Analysis by Recession Separation

	count	mean	std	min	25%	50%	75%	max	tstat	sharpe ratio	MaxDD	α	p_value (α)	β_Market	p_value (β_Market)	R-squared
1974-01 ~ 1975-03	15	0.0413	0.0839	-0.0713	0.0020	0.0278	0.0468	0.2455	1.9061	1.7049	-0.1318	0.0439	0.0687	0.2437	0.4175	0.0512
1975-04 ~ 1979-12	57	0.0190	0.0240	-0.0357	0.0061	0.0163	0.0288	0.0999	5.9847	2.7460	-0.0357	0.0180	0.0000	0.1763	0.0208	0.0933
1980-01 ~ 1980-07	7	0.0308	0.0260	0.0000	0.0119	0.0310	0.0478	0.0651	3.1289	4.0967	0.0000	0.0333	0.0189	-0.1780	0.2873	0.2209
1980-08 ~ 1981-06	11	0.0058	0.0176	-0.0291	-0.0003	0.0090	0.0168	0.0294	1.1023	1.1513	-0.0291	0.0053	0.3383	0.1491	0.2973	0.1197
1981-07 ~ 1982-11	17	0.0204	0.0212	-0.0194	0.0066	0.0190	0.0320	0.0657	3.9605	3.3275	-0.0194	0.0206	0.0007	0.1558	0.0913	0.1783
1982-12 ~ 1990-06	91	0.0138	0.0179	-0.0254	0.0026	0.0113	0.0220	0.0662	7.3385	2.6649	-0.0274	0.0132	0.0000	0.0831	0.0397	0.0467
1990-07 ~ 1991-03	9	0.0214	0.0204	-0.0082	0.0116	0.0211	0.0267	0.0654	3.1476	3.6345	-0.0082	0.0205	0.0042	0.2672	0.0233	0.5443
1991-04 ~ 2001-02	119	0.0165	0.0284	-0.0896	0.0031	0.0147	0.0256	0.1797	6.3316	2.0106	-0.0896	0.0153	0.0000	0.1404	0.0284	0.0404
2001-03 ~ 2001-11	9	0.0190	0.0473	-0.0386	-0.0174	0.0129	0.0518	0.1066	1.2074	1.3942	-0.0386	0.0222	0.1835	0.3527	0.2037	0.2192
2001-12 ~ 2007-11	72	0.0097	0.0178	-0.0241	-0.0036	0.0104	0.0189	0.0593	4.6186	1.8855	-0.0455	0.0097	0.0000	0.0081	0.8947	0.0003
2007-12 ~ 2009-06	19	0.0157	0.0416	-0.0536	-0.0076	0.0127	0.0362	0.1207	1.6446	1.3070	-0.0923	0.0172	0.1074	0.0737	0.5921	0.0172
2009-07 ~ 2020-01	127	0.0066	0.0154	-0.0314	-0.0035	0.0062	0.0164	0.0499	4.8120	1.4792	-0.0648	0.0050	0.0003	0.1451	0.0001	0.1204
2020-02 ~ 2020-06	5	0.0029	0.0421	-0.0323	-0.0201	-0.0112	0.0037	0.0744	0.1549	0.2399	-0.0430	0.0034	0.8652	0.1991	0.3770	0.2629
2020-07 ~ 2023-11	41	0.0030	0.0215	-0.0410	-0.0124	0.0031	0.0176	0.0497	0.8819	0.4771	-0.0740	0.0020	0.5551	0.1070	0.1011	0.0674

Tables & Visualization



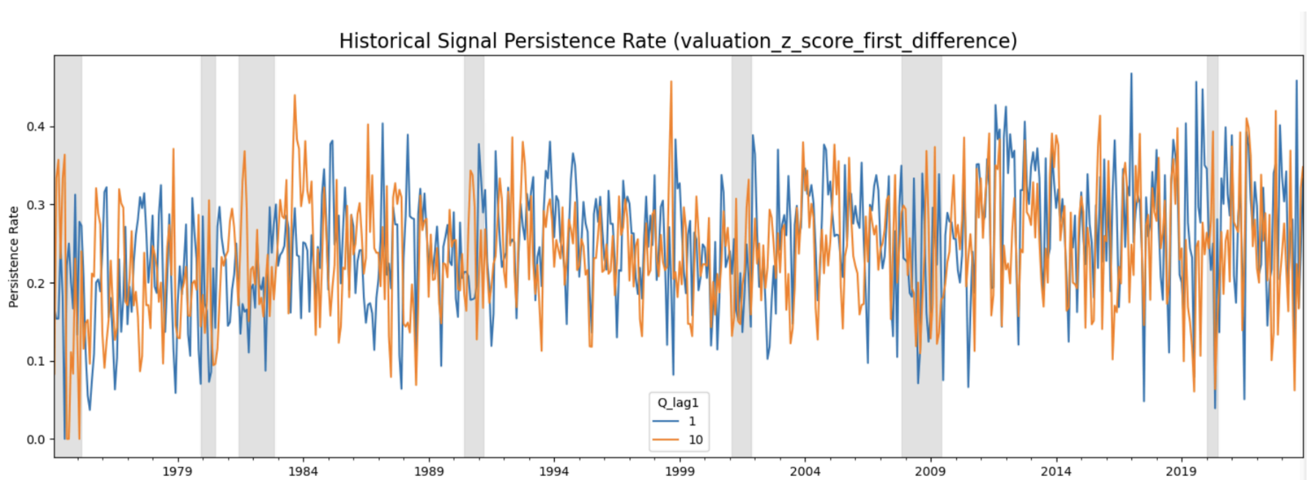


Note:
The "Ranking Persistence Rate" measures the percentage of rankings that remain the same compared to the previous ranking. In other words, it shows how often a ranking position is unchanged from its previous position.

Note:
The Expected Signal Persistent rate is calculate through transition matrix(1month).

The first graph shows the percentage of rankings that remain consistent over the calendar months on an average for each decile stock. More specifically, it depicts how consistent are the ranks of the stocks across the months on average across the years. We can see that top and bottom most deciles are the most consistent, not very, but fairly more than the rest. There is a surprising drop in 6 and/or 7 month. In general, there is not a lot of consistency among the stock for the valuation bias factor.

The graph on the right depicts the probability of a stock holding its rank over the months, and it is clear that apart from the top 2 decile stocks, the drop in probability is drastically high. The persistence probability drops to 0.1 in the first month itself.



Note:
The "Ranking Persistence Rate" measures the percentage of rankings that remain the same compared to the previous ranking. In other words, it shows how often a ranking position is unchanged from its previous position.

Mean-Variance Optimization Performance (without NN)

In addition to the primary factors of FOMO/Familiarity Bias and Valuation Bias, our optimization model incorporates several basic but crucial factors to enhance the robustness and profitability of our trading strategy. These include:

- **Loss Aversion**: This factor exploits investors' tendency to avoid losses more strongly than equivalent gains, allowing us to identify stocks that are likely to experience price corrections.
- **Momentum**: By capitalizing on the continuation of existing price trends, the momentum factor identifies stocks expected to perform well in the near term based on their recent performance.
- **Volatility**: This factor measures the price variability of stocks, enabling the strategy to adjust for risk and target stocks that offer a favorable risk-reward profile.
- **Profitability**: Incorporating fundamental metrics such as return on equity (ROE) and profit margins, the profitability factor identifies financially healthy companies that are expected to generate sustainable returns.
- **Herding Bias**: By analyzing patterns in trading volumes and price movements, the herding bias factor seeks to exploit the tendency of investors to follow the crowd, which can lead to predictable price movements and opportunities for profit.

By integrating these additional factors, our model aims to create a diversified and balanced portfolio that leverages various market inefficiencies and behavioral biases to achieve consistent and superior returns.

The **Mean-Variance Optimization (MVO)** combined strategy demonstrates significant potential in generating consistent and superior returns by integrating various behavioral and technical factors. This approach constructs an investment portfolio that maximizes returns while adhering to specified constraints on risk and leverage. The key performance metrics and analysis based on the results are shown below:

- Mean Returns: The overall strategy achieved a mean return of 0.012041, outperforming the market mean return of 0.006622.
- Standard Deviation: The strategy's standard deviation of 0.026021, lower than the market volatility of 0.045200.
- Sharpe Ratio: The strategy's Sharpe ratio of 1.603052 is significantly higher than the market's Sharpe ratio of 0.507504, indicating a superior return per unit of risk.
- Alpha: While diminishing over time, alpha values remaining positive indicate the strategy's ability to generate returns above the expected market return.

Implemented with a lookback period of 60 months and rebalancing quarterly, the backtesting framework calculates rolling mean returns and covariance matrix for each factor. The MVO is applied at each rebalancing period to generate optimal weights, which are used to calculate portfolio weighted returns.

Neural Networks Ranking:

Neural Networks Ranking: This factor uses 12 features, including 8 technical indicators ('Bollinger Band Percent', 'RSI', 'MACD', 'MACDh', 'MACDs', 'Momentum', 'OBV', 'ATR') and 4 additional metrics ('Size Ranking', 'lagged VWRETX', 'lagged EWRETX', 'lagged S&P 500 return'), each as 30-day time series inputs. These inputs feed into an LSTM model, with the future 30-day return as the label, sorted into cross-sectional deciles: higher returns in higher deciles, lower returns in lower deciles.

The LSTM model has 6 layers, with 4 hidden layers. The first layer has 10 units with a 0.2 a dropout rate, followed by 4 hidden LSTM layer with 12 units and a 0.2 dropout rate. The final dense layer with 10 units uses softmax to rank stocks in 10 deciles. After training on the the previous year's data, we predict daily rankings for this year, short Q1 (bottom 20%) and long Q10 (top 20%) deciles.

Methodology:

- Technical Analysis Indicators: We generate 8 TA indicators ('Bollinger Band Percent', 'RSI', 'MACD', 'MACDh', 'MACDs', 'Momentum', 'OBV', 'ATR') using a Python library. These indicators are preprocessed for stationarity by dividing by price or rolling volume. We also include 'Size', 'lagged VWRETX', 'lagged EWRETX', and 'lagged SPRTRN' as inputs.
- Classify and Label the Return: We use the future 30-day return as our label, calculating cross-sectional deciles and assigning returns to the corresponding deciles: higher returns to higher deciles, lower returns to lower deciles.
- Train the LSTM on Previous Year Data: We input training data and labels into the LSTM, dividing the previous year's data into training and validation sets, and train it year by year.
- Prediction Based on Previous Year Model: We use the next year's feature data to predict each stock's ranking probability for the following year.
- Ranking and Classification: Stocks are ranked based on their highest predicted probability. Q5 (top 20%) has the highest predicted returns, and Q1 (bottom 20%) has the lowest.
- Trading Signals: The strategy longs on Q5 stocks (top 20%) and shorts Q1 (bottom 20%).

Results:

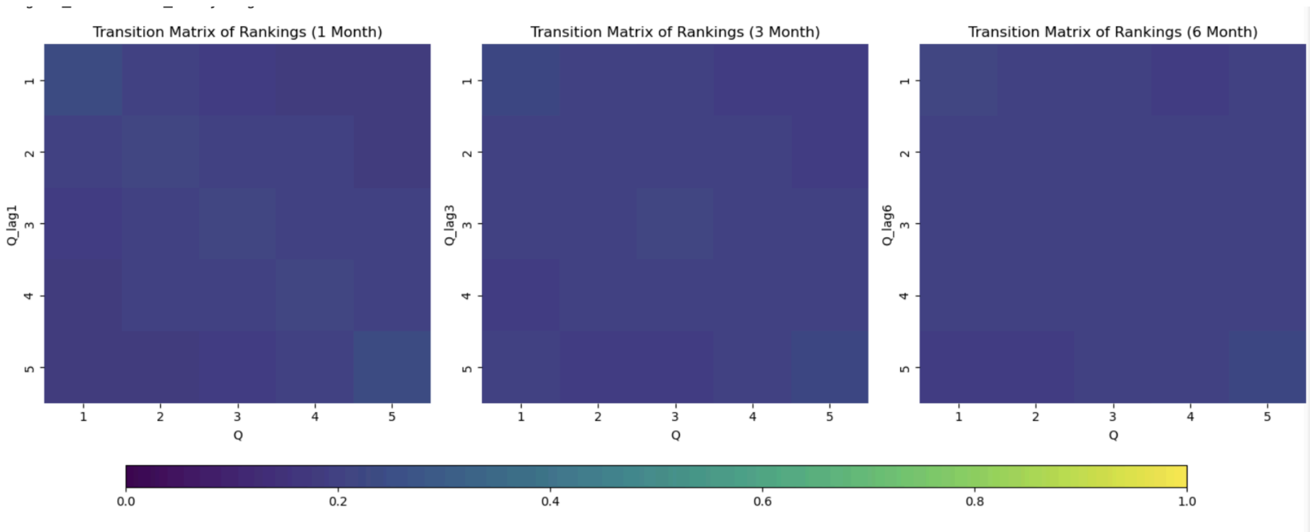
- Whole Periods Performance:
    - The Neural Networks Ranking method shows robust monthly performance across different quantiles.
    - The factor portfolio's mean monthly return is 0.7878% with a Sharpe ratio of 0.5095.
    - The method performs well during normal periods but shows vulnerability during economic recessions, as evidenced by significant dips in the graph during these times.
- 2002-01 to 2007-01:
    - The portfolio shows a mean monthly return of 3.6509% with a Sharpe ratio of 4.7664.
    - Quantiles show varied performance, with the Q5 portfolio having the highest returns and lower standard deviation.
- 2010-01 to 2020-01:
    - The portfolio has a mean monthly return of 1.5659% with a Sharpe ratio of 1.4324.

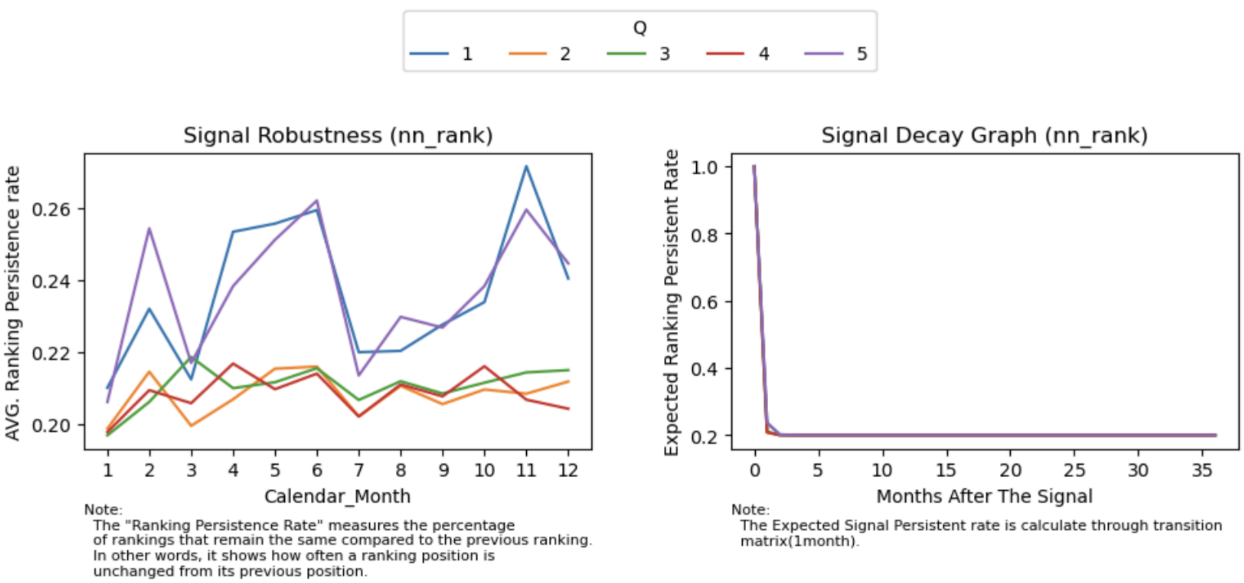
    - The Q5 portfolio demonstrates the highest returns with relatively low standard deviation, while the Q1 portfolio has the lowest returns.
- Overall Insight:
    - The Neural Networks Ranking method performs effectively during normal periods.
    - However, it tends to underperform during economic recessions, highlighting its sensitivity to market downturns.
- Possible improvement:
    - Right now we are training a one-year model with previous year data and predict on stock quintile/decile based on current year lagged indicators. This works fine during normal years but when economic recessions or crises come, our model will not be updated in time. So one possible improvement would be to update our model quarterly/monthly to reduce the impact of economic recessions or crises and thus avoid those downsides lost in our model's prediction.

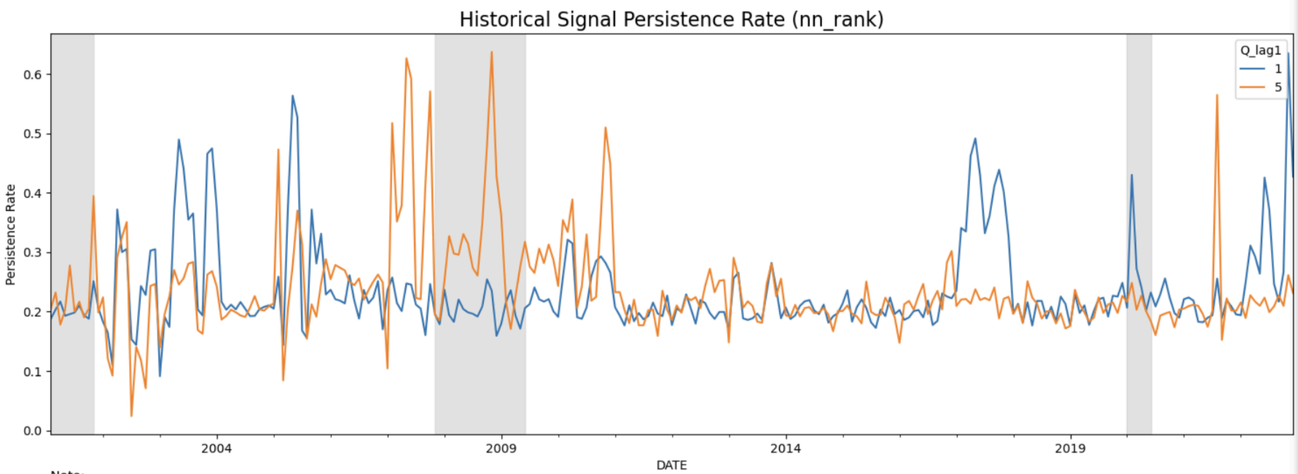
Alpha Analysis by Recession Separation

	count	mean	std	min	25%	50%	75%	max	tstat	sharpe ratio	MaxDD	α	p_value (α)	β_Market	p_value (β_Market)	R-squared
2001-01 ~ 2001-02	2	0.0351	0.0530	-0.0024	0.0163	0.0351	0.0538	0.0726	0.9354	2.2913	-0.0024	0.0537	nan	0.5439	nan	1.0000
2001-03 ~ 2001-11	9	-0.0396	0.0495	-0.1271	-0.0637	-0.0259	0.0007	0.0194	-2.3977	-2.7687	-0.2142	-0.0395	0.0627	0.0128	0.9670	0.0003
2001-12 ~ 2007-11	72	0.0439	0.0546	-0.1063	0.0022	0.0545	0.0857	0.1421	6.8226	2.7853	-0.4096	0.0450	0.0000	-0.2131	0.2541	0.0185
2007-12 ~ 2009-06	19	-0.0819	0.0724	-0.2612	-0.1202	-0.0542	-0.0324	-0.0070	-4.9291	-3.9173	-0.8079	-0.0885	0.0001	-0.3269	0.1609	0.1122
2009-07 ~ 2020-01	127	0.0100	0.0570	-0.1671	-0.0095	0.0050	0.0528	0.1668	1.9800	0.6086	-0.5291	0.0125	0.0184	-0.2285	0.0974	0.0218
2020-02 ~ 2020-06	5	0.0415	0.0236	0.0132	0.0276	0.0404	0.0513	0.0751	3.9379	6.1005	0.0000	0.0411	0.0246	-0.1293	0.2903	0.3535
2020-07 ~ 2022-12	30	-0.0041	0.0262	-0.0550	-0.0224	-0.0047	0.0110	0.0502	-0.8662	-0.5479	-0.2579	-0.0048	0.3343	0.0776	0.3888	0.0266

Signal analysis



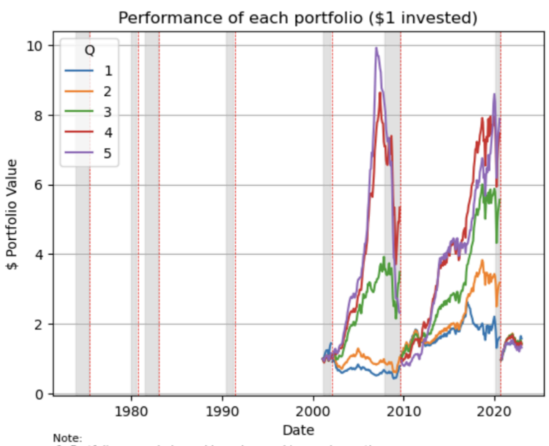




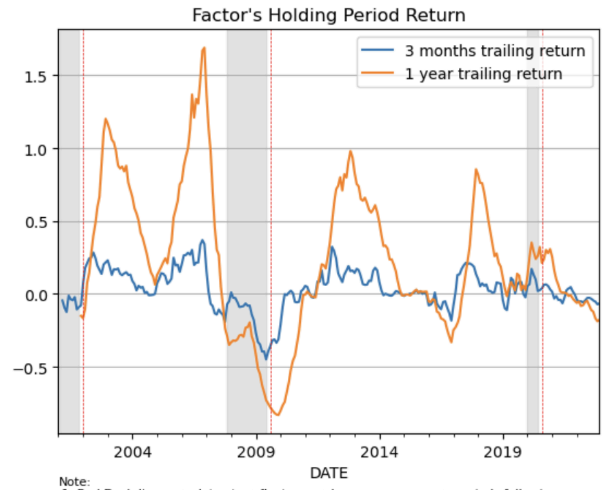
Note:
The "Ranking Persistence Rate" measures the percentage of rankings that remain the same compared to the previous ranking.
In other words, it shows how often a ranking position is unchanged from its previous position.

Whole Periods Performance

```
---------- Performance of Each Quantiles Portfolios----------
   count      mean       std        min       25%       50%       75%       max
Q
1  264.0  0.008299  0.065674 -0.267770 -0.030430  0.009255  0.044927  0.262273
2  264.0  0.010919  0.058116 -0.182440 -0.023213  0.013124  0.044265  0.248626
3  264.0  0.018868  0.059639 -0.182388 -0.012738  0.022785  0.054535  0.248401
4  264.0  0.022054  0.059470 -0.181116 -0.010093  0.026358  0.058334  0.197715
5  264.0  0.018436  0.064688 -0.239654 -0.016929  0.021905  0.061592  0.197453
---------- Performance of Factor Portfolio----------
count     264.000000
mean        0.010137
std         0.062824
min        -0.261191
25%        -0.017213
50%         0.004578
75%         0.057164
max         0.166827
tstat       2.621738
sharpe      0.558956
Name: RET, dtype: float64
```



Note:
1. Portfolios are rebalanced based on ranking each month.
2. Q: Ranked by nn_rank. The higher the Q, the higher the rank.
3. Red Dash lines are reset dates to reflect several economy recovery periods following a recession. At each reset date, all portfolio values are reset to $1 to make subsequent performance comparable.
3. Grey Areas are the recession periods.



Note:
1. Red Dash lines are dates to reflect several economy recovery periods following a recession.
2. Grey Areas are the recession periods.

2002-01 ~ 2007-01

---------- Performance of Each Quantiles Portfolios----------

Q	count	mean	std	min	25%	50%	75%	max
1	60.0	-0.010048	0.042286	-0.151271	-0.034230	-0.007083	0.019128	0.073697
2	60.0	-0.002882	0.043884	-0.103085	-0.031378	-0.009104	0.026499	0.086143
3	60.0	0.021098	0.043248	-0.105064	-0.008597	0.024480	0.052428	0.104000
4	60.0	0.035522	0.035336	-0.053831	0.008466	0.039160	0.056723	0.111548
5	60.0	0.040746	0.041181	-0.093214	0.011828	0.041694	0.071414	0.120996

---------- Performance of Factor Portfolio----------

count	60.000000
mean	0.050794
std	0.035611
min	-0.018198
25%	0.027379
50%	0.049721
75%	0.079225
max	0.121672
tstat	11.048574
sharpe	4.941072

Name: RET, dtype: float64

2010-01 ~ 2020-01

---------- Performance of Each Quantiles Portfolios----------

Q	count	mean	std	min	25%	50%	75%	max
1	120.0	0.005610	0.048544	-0.144602	-0.025756	0.004652	0.036402	0.143780
2	120.0	0.008956	0.042527	-0.135201	-0.014750	0.010632	0.033824	0.140291
3	120.0	0.014314	0.044029	-0.107997	-0.006499	0.018519	0.040513	0.147022
4	120.0	0.018523	0.047029	-0.102501	-0.003733	0.019221	0.046839	0.146964
5	120.0	0.020893	0.045250	-0.097833	-0.002390	0.022178	0.048270	0.133803

---------- Performance of Factor Portfolio----------

count	120.000000
mean	0.015283
std	0.037244
min	-0.078307
25%	-0.004578
50%	0.008242
75%	0.043109
max	0.109946
tstat	4.494980
sharpe	1.421438

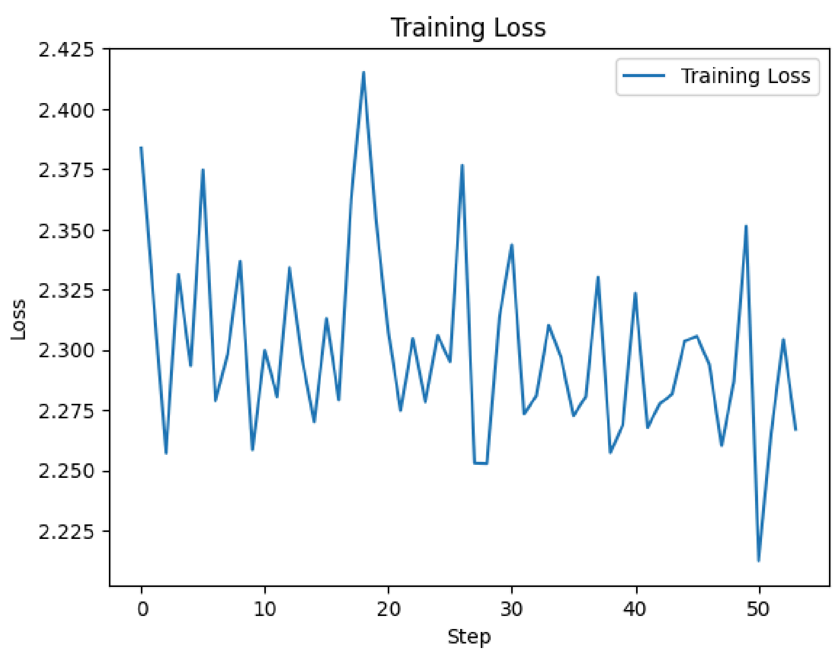
Name: RET, dtype: float64

CQT + Convolution Neural Network (CNN) Classifier:

This strategy uses daily returns in a window of the past 30 days to predict the monthly return of the next month. The past 30 days daily returns are passed into the CQT model to get the transformed output with 20 features for each of the past 30 days. This output is passed through a CNN classifier which consists of a fully connected layer (FCN) consisting of 10 neurons which is essentially 10 deciles to be classified. We modeled this strategy as a multi-class classification problem.

Unfortunately, the model does not learn anything contrary to our hypothesis. Below is a plot of the training loss for 1 epoch across various batches. The training loss is noisy and it continues to be in the same range across epochs. The model not learning anything was reinforced by the predictions it made on the evaluation set after training. It consistently produced identical logits for any stock data input into the model. There can be multiple reasons for the failure of this approach. One can be that CQT does not capture the kind of information that can help in prediction of future monthly returns given the daily returns of the past 30 days. If we have more granular data or high frequency data for that timeline, it might help us but the daily returns are not helpful.

Another pitfall can be the hyperparameters used for training the model. We used a kernel of size 30x20 to account for all the features of the 30 days at the same time and even tried a kernel of 3x3 but none of them seem to work well.



Viability Concerns

While this trading strategy is clearly displaying signs of success, it is important to acknowledge that this strategy may have some key viability concerns in its implementation versus this strategy assessment. First, this strategy relies on various data sources, including trading volumes, historical financial data, and new financial data. The quality and consistency of this data can significantly impact the performance of this strategy, as inaccuracies and limitations in data can lead to misleading signals or suboptimal portfolio construction. As this strategy involves frequent trading, implementing this strategy may incur transaction costs that affect returns or be hindered by real-time challenges like data delays and managing execution risks. Lastly, this strategy's effectiveness relies on exploiting market inefficiencies and behavioral biases. As the market becomes more efficient, these biases may diminish, in turn reducing the strategy's alpha and profitability.

Robustness

This strategy incorporates a combination of factors, including behavioral biases, valuation metrics, and technical analysis. This diversification helps mitigate the risk of relying on a single factor that may become less effective over time. By employing the MVO, the strategy optimizes the portfolio for both return and risk, helping to construct a balanced portfolio that can withstand market fluctuations. This portfolio must be periodically rebalanced so weights of each factor can be adjusted based on the changing market conditions. This helps to maintain the desired risk profile while capturing new opportunities. In comparison, machine learning techniques, like neural networks, allow the ranking signal to adapt to evolving market dynamics, showing its robustness, its ability to continuously learn from new data and potentially identify new trading opportunities. However, predicting recessions becomes integral to the NN model.

Behavioral Explanation

Our strategy's success hinges on its ability to exploit psychological biases. By understanding and anticipating these biases, we can combine those with machine learning model to build a better rating system than MVO, thereby achieving consistent returns. Below we have detailed the behavioral phenomena underlying two of our key factors that we have developed:

Familiarity bias occurs when investors prefer familiar investments, such as well-known companies or local stocks, often due to a perceived reduction in risk. This can lead to overinvestment in familiar assets and underinvestment in unfamiliar ones. Fear of Missing Out (FOMO) drives investors to buy into trends and popular stocks to avoid missing potential gains, often amplifying market movements and leading to positive feedback training.

Valuation bias exploits the tendency of markets to misprice assets relative to their ”intrinsic valuation”. Investors often overreact/underreact to recent performance or news, leading to deviations from intrinsic value. By focusing on drastic change in valuation metrics like earnings-to-price (E/P), the strategy identifies undervalued stocks expected to revert to their mean valuation, capturing gains from the correction of these mispricings.

CODE: https://github.com/tushar2407/MGTF430