Final Project: An investment strategy based on firms' intangible assets

Code for the Chicago Booth course on Quantitative Portfolio Management

Group Members: Charene Shen, Arthur Ji, Anthony Song, Amy Cai

Strategy Overview

In our technology sector strategy, we develop a signal that assigns lower ratings to firms with high intangible assets on their balance sheets. We integrate intangibility score with a value signal, acknowledging how the market might have already factored in these risks.

We then test our strategy and benchmark it against fama-french 5 factor industry performance in the technology sector.

Economic Idea

We think high intangible assets ratio is concerning due to several key reasons:

1. Valuation Challenges: Intangible assets like patents and trademarks are difficult to value precisely, leading to possible overvaluation risks.
2. Impairment Risks: These assets are prone to value reductions from market or performance shifts, impacting financial health.
3. Potential for Manipulation: Elevated intangible asset levels may indicate aggressive accounting, with inflated asset values to enhance financial appearance.
4. Liquidity Concerns: During financial distress, intangible assets are less liquid than tangible assets, posing challenges in quick cash conversion for meeting financial obligations.

The 2001 AOL Time Warner merger exemplifies these risks. Following the merger, the company grappled with inflated goodwill. The dot-com bubble burst revealed these overvaluations, leading to a $99 billion goodwill impairment in 2002 and a subsequent sharp stock price decline. This case highlights the importance of careful evaluation of intangible asset valuations.

We also integrate intangibility ratio with a value signal in our strategy, recognizing that the market may have already adjusted for these risks. This combination allows us to balance the inherent uncertainties of high intangible asset ratios with the market's perception of value. We will cover details for signal construction in later sections.

In [1]:
```
import qpm_download_modified
import qpm_modified
import qpm
import pandas as pd
import numpy as np
import wrds
import statsmodels.api as sm
from datetime import datetime
from dateutil.relativedelta import relativedelta
import requests
from bs4 import BeautifulSoup
from scipy.stats import zscore
from pandas_datareader import data as web
import matplotlib.pyplot as plt
import matplotlib.pylab as pylab
import matplotlib.dates as mdates


import warnings
warnings.simplefilter(action='ignore', category=FutureWarning)
```

Data Overview

Data Reporting and Lag Consideration:

We account for a typical 3-month lag in the reporting of quarterly financials. For example, the fundamentals data for Q1 of the fiscal year 2020 is assumed to be available by June 2020. This lag reflects the time companies take to finalize and report their quarterly results by the end of each quarter.

Data Sourcse:

The strategy primarily relies on data sourced from comprehensive financial databases: Compustat.fundq, Crsp, Ff.fivefactors_monthly.

Data Handling and Processing:

To ensure the timeliness and relevance of the data, the strategy incorporates the most recent quarterly data available, considering the 3-month reporting lag. In cases where the latest quarterly data is unavailable, the strategy employs a method to 'fill in' missing data. For any missing data point, we use the last available data within the one fiscal year. If data is still missing (i.e., not reported within the year), those data points are excluded from the analysis.

Key metrics:

Key metrics include intangible assets, total assets, total current assets, industry classification, etc..

In [2]:
```
# Data Preparation
start_date = '2010-01-01'
end_date = '2023-08-31'
fundq_variables = ['intanq', 'gdwlq', 'atq', 'actq', 'rdq', 'mkvaltq', 'dlttq', 'dlcq', 'ltq']
```

In [3]:
```
df_full = qpm_download_modified.prepare_data(start_date, end_date, fundq_variables)

# fill unavailable data with the last available data within one year
for column in fundq_variables:
    print(f"Processing column: {column}")
    df_full[column] = df_full.groupby(['permno'], sort=False)[column].apply(lambda x: x.ffill(limit = 12))
print("All null data processed.")
# drop other unavailable data
df_full.dropna(inplace = True)
df_full = df_full.groupby(['permno','ldate']).last().reset_index()
```

```
Enter your WRDS username [tianqishen]:charene
Enter your password:········
WRDS recommends setting up a .pgpass file.
Create .pgpass file now [y/n]?: y
Created .pgpass file successfully.
You can create this file yourself at any time with the create_pgpass_file() function.
Loading library list...
Done
Processing column: intanq
Processing column: gdwlq
Processing column: atq
Processing column: actq
Processing column: rdq
Processing column: mkvaltq
Processing column: dlttq
Processing column: dlcq
Processing column: ltq
All null data processed.
```

In [4]:
```
# include NAICS data to filter stocks by industry
IMPORT_NAICS_DATA = True

if IMPORT_NAICS_DATA:
    # Retrieve NAICS code information from guidance for economic consensus data
    url = "https://www.census.gov/programs-surveys/economic-census/year/2022/guidance/understanding-naics.html"
    response = requests.get(url)
    soup = BeautifulSoup(response.content, 'html.parser')
    tables = soup.find_all('table')
    df_naics = pd.read_html(str(tables[1]))[0]
    df_naics = pd.DataFrame(df_naics)
    def transform_sector(row):
        if '-' in row['Sector']:
            int1, int2 = map(int, row['Sector'].split('-'))
            return pd.DataFrame([row] * (int2 - int1 + 1)).assign(Sector=range(int1, int2 + 1))
        else:
            return pd.DataFrame([row])

    df_naics = pd.concat([transform_sector(row) for _, row in df_naics.iterrows()], ignore_index=True)
    df_naics['Sector'] = df_naics['Sector'].astype(str)
# include a sector column
df_full['naics sector'] = df_full['naics'].astype(str).str[:2]
sector_to_description = df_naics.set_index('Sector')['Description'].to_dict()
df_full['naics sector'] = df_full['naics sector'].map(sector_to_description)

# include sic data
file_name = 'SICS_CODE.xlsx'
df_sic = pd.read_excel(file_name)
sics_to_office = df_sic.set_index('SIC Code')['Office'].to_dict()
df_full['sic'] = df_full['sic'].astype(int)
df_full['sic office'] = df_full['sic'].map(sics_to_office)
```

Step 1. Construct Signal

Signal Construction for 'Intan' Strategy

Signal: The strategy focuses on the ratio of intangible assets (intanq) to net assets (total assets atq minus current assets actq), **negated**. This signal assesses the proportion of intangible assets in a company's asset structure. A higher ratio may indicate a significant reliance on intangible assets, which can vary in value and stability compared to tangible assets.

Signal Construction for 'Intan_value' Strategy

The 'Intan_value' strategy combines signals from intangible assets and the book-to-market ratio, calculating their z-scores within each sector to account for industry-specific norms regarding intangible assets. The final signal is derived by averaging these z-scores for each signal.

In [5]:
```
_REMOVE_MICRO_CAPS = False
_INDUSTRY_REFERENCE = 'sic office' # choose between 'sic office' or 'naics sector'
_INDUSTRY_CHOSEN = 'Office of Technology'

# create signal for first strategy
df_full['intan_to_noncurrentasset_neg'] = - df_full['intanq'] / (df_full['atq'] - df_full['actq'])
# create signal for the second strategy
```

```
df_full['bv'] = df_full['atq'] - df_full['ltq']
df_full['bvtm'] = df_full['bv'] / df_full['mkvaltq']
df_full['zscore_1'] = df_full.groupby(['ldate',_INDUSTRY_REFERENCE])['intan_to_noncurrentasset_neg'].transform(lambda
df_full['zscore_2'] = df_full.groupby(['ldate',_INDUSTRY_REFERENCE])['bvtm'].transform(lambda x: zscore(x, nan_policy=
df_full['zscore'] = (df_full['zscore_1'] + df_full['zscore_2']) / 2
```

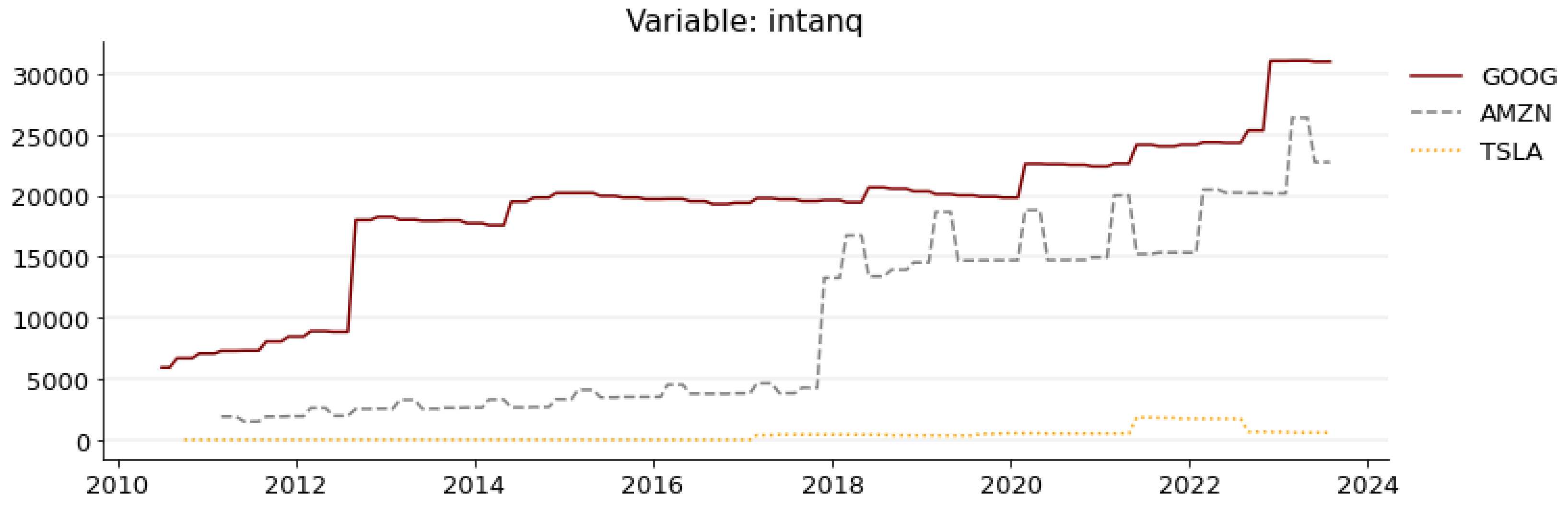
Case Study of Key Variables

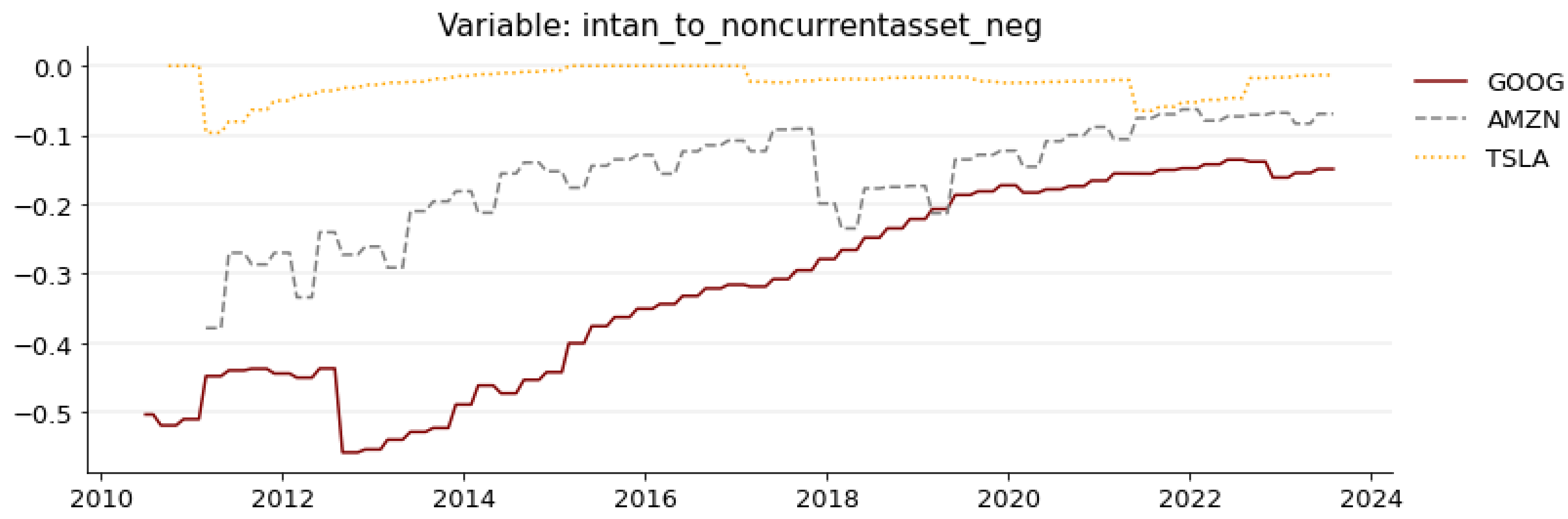
Below are key variables and signals for Google (GOOG), Amazon (AMZN), and Tesla (TSLA).

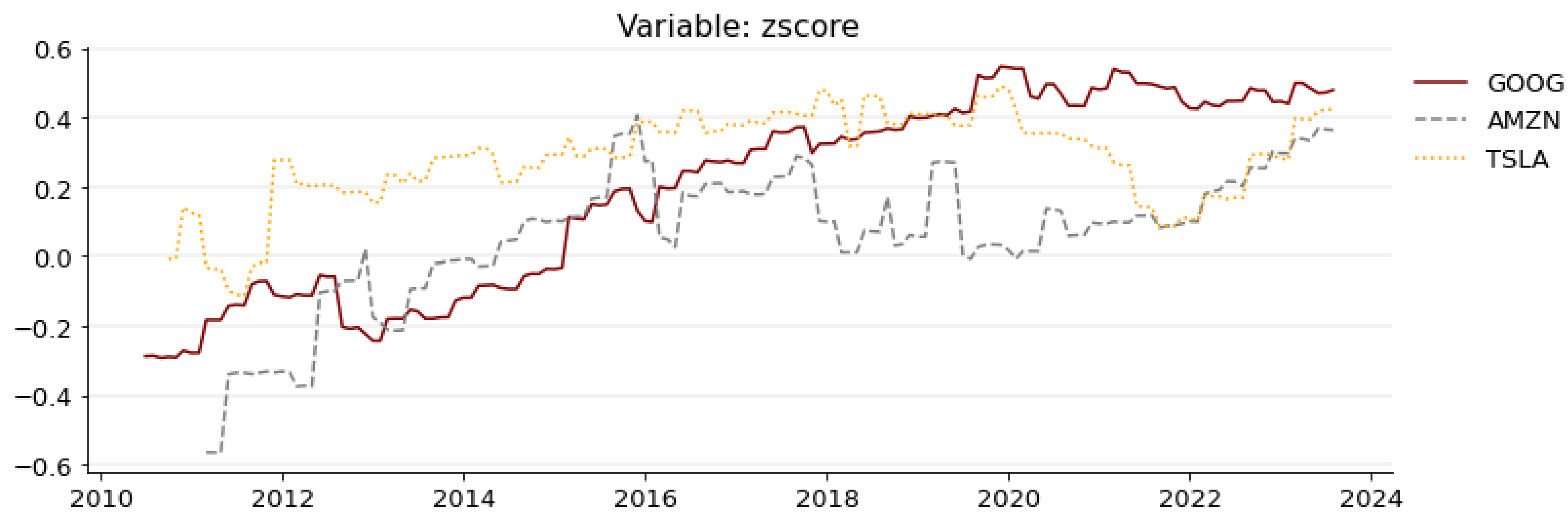
The graphs shows that there could be major changes to intangible assets quarter over quarter. The major reasons for a change in a company's intangible assets typically include:

1. Acquisitions: Buying companies at a substantial premium.
2. Internal Development: Investing in creating new intangible assets (e.g., R&D, brand development
3. Amortization and impairment losses of intangible assets like patents and copyrights.

In [6]:
```
qpm.plot_variables(df_full, variable_list = ['intanq'],  id_type = 'ticker', id_list = ['GOOG', 'AMZN', 'TSLA'],
                   start_date = start_date, end_date = end_date)
qpm.plot_variables(df_full, variable_list = ['intan_to_noncurrentasset_neg'],  id_type = 'ticker', id_list = ['GOOG',
                   start_date = start_date, end_date = end_date)
qpm.plot_variables(df_full, variable_list = ['zscore'],  id_type = 'ticker', id_list = ['GOOG', 'AMZN', 'TSLA'],
                   start_date = start_date, end_date = end_date)
df_sum = df_full.sort_values(['ldate','ticker'])
print(df_sum[['ldate','ticker','me','daret','intan_to_noncurrentasset_neg', 'zscore']].loc[df_sum['ticker'].isin(['AAI
```








```
             ldate ticker            me     daret  \
521344 2023-08-01   TSLA  8.191443e+05 -0.034962
360921 2023-09-01   AMZN  1.313150e+06 -0.078907
521345 2023-09-01   TSLA  7.954494e+05 -0.030456

        intan_to_noncurrentasset_neg    zscore
521344                     -0.013300  0.422534
360921                     -0.067586  0.416157
521345                     -0.013892  0.401426
```

Sample Selection

We do not lag signal because a three month report lag has been assumed.

In [7]:
```
_STRATEGY_NAME = 'Intan'
if _STRATEGY_NAME == 'Intan':
    df_full['signal'] = df_full['intan_to_noncurrentasset_neg']

elif _STRATEGY_NAME == 'Intan_value':
    df_full['signal'] = df_full['zscore']
else:
    raise Exception('Please provide a valid _STRATEGY_NAME..')

df_sum = df_full.sort_values(['ldate','ticker'])
```

In [8]:
```
# Select the relevant sample
df = qpm.select_sample(df_full, sample_start = start_date, sample_end = end_date,
                       remove_micro_caps = _REMOVE_MICRO_CAPS)
print('> Select Industry for Analysis...')
df_industry = df_full[df_full[_INDUSTRY_REFERENCE]==_INDUSTRY_CHOSEN]
```

```
> Selecting Sample for Given Criteria...
> Select Industry for Analysis...
```

Step 2. Portfolio Construction

Next, we sort the stocks into portfolios:

- retP_rank_longonly: Rank-based long-only portfolio
- retP_rank_longshort: Rank-based long-short portfolio
- retP_vw_P1, ..., retP_vw_P5: The returns on the 5 portfolios sorted by the signal (value or size) and weighted by market capitalization
- retF_vw: The return on the factor, which is retP_vw_P5-retP_vw_P1

In [9]:
```
_SORT_FREQUENCY = 'Quarterly'
_NUM_PORT = 5

# portfolio contruction for 'Intan_value'
_STRATEGY_NAME = 'Intan_value'
print('> Portfolio construction for ' + _STRATEGY_NAME)
if _STRATEGY_NAME == 'Intan':
    df_full['signal'] = df_full['intan_to_noncurrentasset_neg']

elif _STRATEGY_NAME == 'Intan_value':
    df_full['signal'] = df_full['zscore']
else:
    raise Exception('Please provide a valid _STRATEGY_NAME..')

df_sum = df_full.sort_values(['ldate','ticker'])
# Select the relevant sample
df = qpm.select_sample(df_full, sample_start = start_date, sample_end = end_date,
                       remove_micro_caps = _REMOVE_MICRO_CAPS)
print('> Select Industry for Analysis...')
df_industry = df_full[df_full[_INDUSTRY_REFERENCE]==_INDUSTRY_CHOSEN]
df_industry2, df_rets2 = qpm_modified.create_portfolios_modified(df_industry, sort_frequency = _SORT_FREQUENCY, num_po
print(df_rets2.tail())
```

```
> Portfolio construction for Intan_value
> Selecting Sample for Given Criteria...
> Select Industry for Analysis...
> Sorting stocks into 5 portfolios at frequency: Quarterly...
> Computing returns using various weights...
          ldate  retP_rank_longonly  retP_rank_longshort  retP_vw_P1  \
160 2023-05-01            0.040319             0.029211    0.001131   
161 2023-06-01            0.077054             0.025378    0.069838   
162 2023-07-01            0.069515             0.040823    0.069069   
163 2023-08-01           -0.067203             0.002791   -0.019102   
164 2023-09-01           -0.088077            -0.009896   -0.052957   

     retP_vw_P2  retP_vw_P3  retP_vw_P4  retP_vw_P5   retF_vw
160    0.039179    0.012315    0.064948    0.117610  0.116478
161    0.069811    0.068157    0.049395    0.043082 -0.026756
162    0.060226    0.046271   -0.003138    0.101819  0.032750
163    0.010675   -0.010771   -0.017315   -0.032941 -0.013839
164   -0.065257   -0.037878   -0.040969   -0.031236  0.021721
```

Step 3. Portfolio Analytics

We first plot the average returns on the portfolios. Then, we plot the cumulative returns on various strategies. For the long-only strategy, we use the market as a simple benchmark. For the long-short strategies, we use the risk-free rate as a benchmark. Later, we will use regression analysis to properly correct for the factors.

Strategy 2: 'Intan'

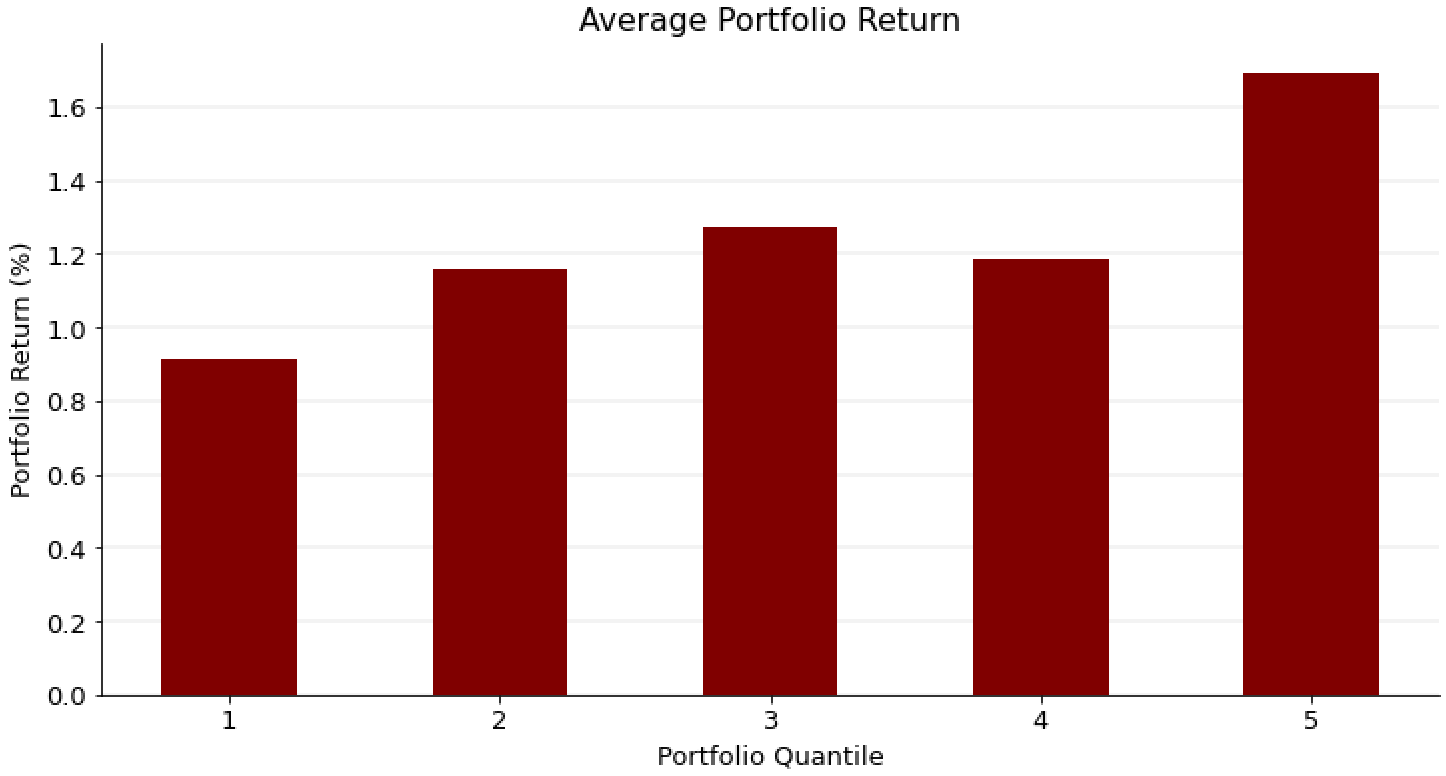
Observations on performances:

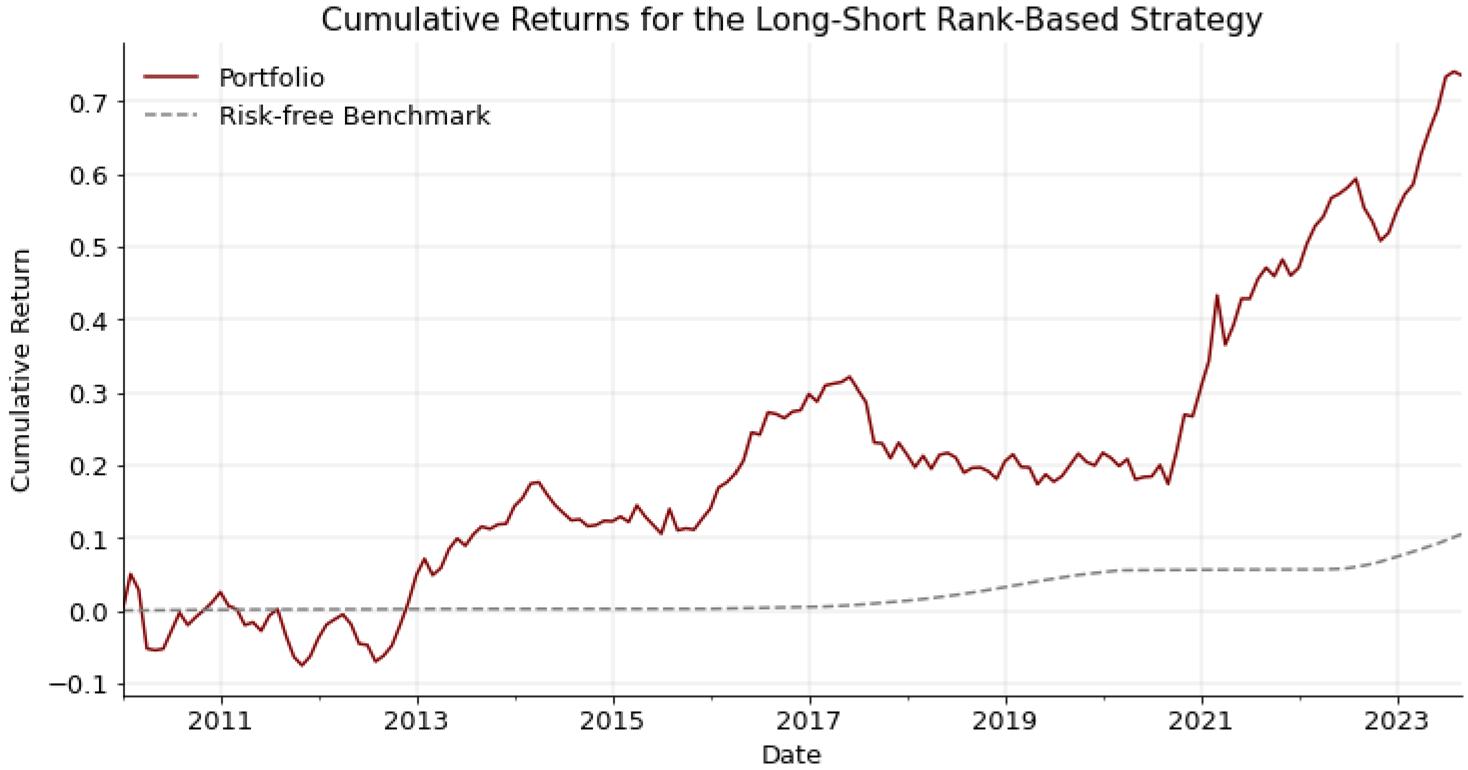
1. Positive relationship between the quintile rank and portfolio performance
2. Outperform risk-free benchmark
3. Periods of outperformance and underperformance compared to FF industry portfolios ##### Observations on alphas:

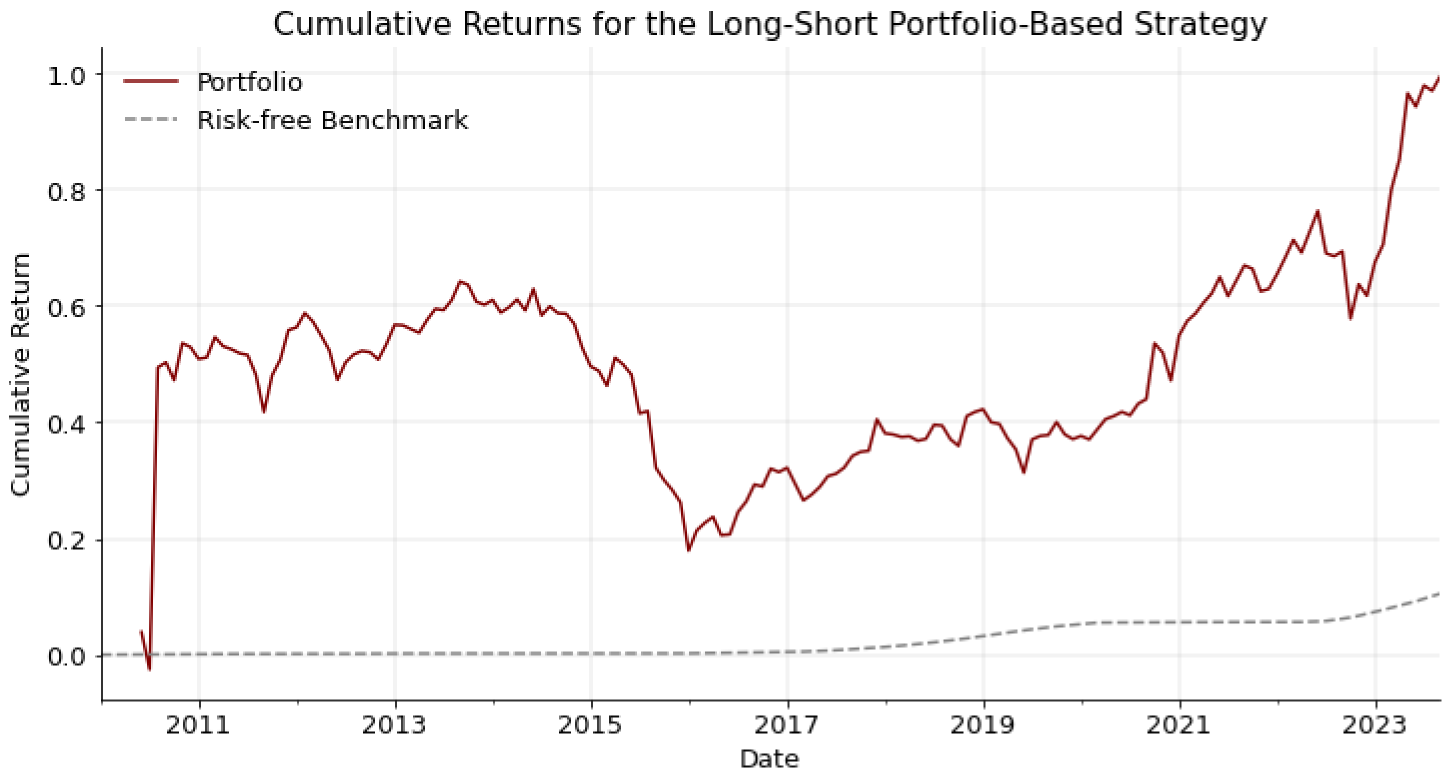
4. As the number of factors increases, the significance of performance disparities escalates, suggesting that adding more factors does not account for, but rather highlights, unexplained performance variations.
5. The most pronounced effects are observed within rank-based and long-short portfolios, which implies these portfolios are more efficient in capitalizing on the signal. This makes sense becausse long-short strategies are aligned with our economic rationale, as they emphasize the negative connotations of high ratios over any positive connotations of low ratios, indicating a preference for short/long-short strategies strategies. ##### Observations on betas: Observations on Betas:
6. In the first table, the betas of the CAPM and the Fama-French Three-Factor Model exhibit higher significance for the rank-based strategies (2), (3), (5), and (6), indicating a stronger relationship with these strategies compared to others.
7. In the second table, the positive and significant Small Minus Big (SMB) coefficients for the rank-based strategies (2), (3), and (5) indicate a preference for smaller companies within these strategies.
8. The consistently negative and significant Up Minus Down (UMD) coefficients suggest that these rank-based strategies do not capitalize on momentum effects.
9. The lack of consistent patterns in the betas across different strategies suggests that these factors provide limited explanatory power regarding the performance of the signals.

In [10]:
```
qpm.analyze_strategy(df_rets2, analysis_type = 'Performance')
```

```
> Merging strategy returns with Fama and French factor returns...
```

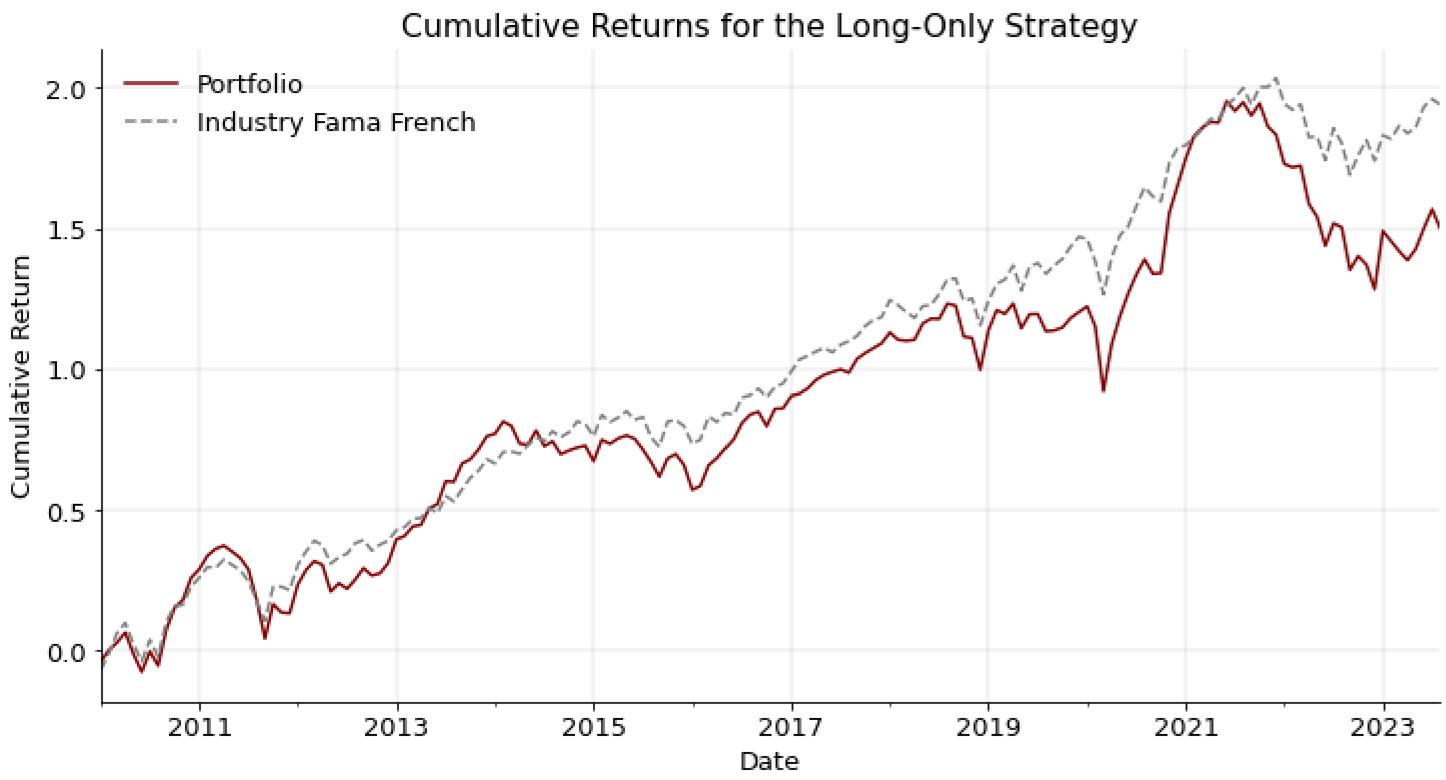






In [11]:

```
# Include fama-french 5 factor industry performance for comparsion
ff_factor = 'F-F_Research_Data_5_Factors_2x3'
ff_factor_data = web.DataReader(ff_factor, 'famafrench', start=start_date, end=end_date)[0]
ff_portfolio = '49_Industry_Portfolios'
ff_portfolio_data = web.DataReader(ff_portfolio, 'famafrench', start=start_date, end=end_date)[0]
ff_portfolio_data = ff_portfolio_data.sub(ff_factor_data.RF, axis=0)/100
sectors = ['Hardw', 'Softw', 'Chips', 'LabEq', 'ElcEq', 'Telcm']
ff_portfolio_data['ff Sector Performance'] = ff_portfolio_data[sectors].mean(axis=1)
ff_portfolio_data = ff_portfolio_data.reset_index().sort_values(['Date']).rename(columns = {'Date' : 'ldate'})
ff_portfolio_data['ldate'] = ff_portfolio_data['ldate'].dt.strftime('%Y-%m') + '-01'
ff_portfolio_data['ldate'] = pd.to_datetime(ff_portfolio_data['ldate'])

df2 = pd.merge(ff_portfolio_data, df_rets2, on = ['ldate'], validate = 'many_to_one', indicator = True).reset_index()
df2.sort_values(['ldate'], inplace = True)
df2['CLNretP'] = np.log(1 + df2['retP_rank_longonly']).cumsum()
plot_df = df2[[ 'CLNretP', 'ff Sector Performance','ldate']].set_index('ldate')
fig = plt.figure(figsize = (12, 6))
plot_df['CLNretP'].plot(ax = plt.gca(), color = 'maroon', label = 'Portfolio')
plot_df['ff Sector Performance'].cumsum().plot(ax = plt.gca(), color = 'gray', label = 'Industry Fama French')
plt.xlabel('Date'); plt.ylabel('Cumulative Return')
plt.title('Cumulative Returns for the Long-Only Strategy')
plt.legend()
plt.show()
plt.close()
```



In [12]:
```
qpm.analyze_strategy(df_rets2, analysis_type = 'Summary')
```

```
> Merging strategy returns with Fama and French factor returns...
                     count      mean       std       min       25%       50%  \
retP_rank_longonly   165.0  0.010550  0.063617 -0.205728 -0.028529  0.013069
retP_rank_longshort  165.0  0.004071  0.022194 -0.077668 -0.010772  0.005313
retF_vw              160.0  0.007135  0.063067 -0.111973 -0.016770  0.002533
mktrf                165.0  0.010054  0.044228 -0.133900 -0.015500  0.012900
smb                  165.0 -0.000310  0.025981 -0.082800 -0.018200  0.001200
hml                  165.0 -0.001476  0.033048 -0.138700 -0.018600 -0.004200

                          75%       max
retP_rank_longonly   0.044227  0.237605
retP_rank_longshort  0.017990  0.094211
retF_vw              0.023001  0.680430
mktrf                0.034400  0.136500
smb                  0.015700  0.071100
hml                  0.013700  0.127500
```

In [13]:
```
qpm.analyze_strategy(df_rets2, analysis_type = 'Factor Regression')
```

```
> Merging strategy returns with Fama and French factor returns...

----------------------------------------------------------------
> Running Factor Regressions: Table 1  - 3 Fama-French Factors
----------------------------------------------------------------

=========================================================================
                (1)      (2)      (3)      (4)      (5)       (6)
-------------------------------------------------------------------------
const          0.0084   -0.0030   0.0042** 0.0086   -0.0007    0.0049***
               (0.0051) (0.0025)  (0.0017) (0.0052) (0.0018)   (0.0017)
mktrf          -0.1208  1.2845*** 0.0257   -0.1435  1.0980***  -0.0054
               (0.1137) (0.0550)  (0.0382) (0.1229) (0.0429)   (0.0401)
hml                                        -0.0900  -0.2133*** 0.1016**
                                           (0.1568) (0.0547)   (0.0512)
smb                                        0.1227   0.9031***  0.1233*
                                           (0.2151) (0.0751)   (0.0702)
R-squared      0.0071   0.7752    0.0029   0.0104   0.8836     0.0602
R-squared Adj. 0.0008   0.7737    -0.0035  -0.0087  0.8814     0.0421
N              160      160       160      160      160        160
R2             0.01     0.78      0.00     0.01     0.88       0.06
=========================================================================
Standard errors in parentheses.
* p<.1, ** p<.05, ***p<.01
(1): Long-Short Value Weights ~ CAPM Model
(2): Long-Only Rank Weights ~ CAPM Model
(3): Long-Short Rank Weights ~ CAPM Model
(4): Long-Short Value Weights ~ 3-Factor Fama French Model
(5): Long-Only Rank Weights ~ 3-Factor Fama French Model
(6): Long-Short Rank Weights ~ 3-Factor Fama French Model
Annualized Information Ratios:
                       1      2      3      4      5      6
Alpha              0.101 -0.036  0.051  0.103 -0.009  0.058
Std(resid)         0.217  0.105  0.073  0.217  0.076  0.071
Information Ratio  0.464 -0.345  0.700  0.474 -0.114  0.826



----------------------------------------------------------------
> Running Factor Regressions: Table 2  - 5 Fama-French Factors + Momentum
----------------------------------------------------------------

============================================================================
                (1)      (2)       (3)      (4)       (5)        (6)
----------------------------------------------------------------------------
const          0.0099*  0.0008     0.0047*** 0.0108** 0.0015     0.0053***
               (0.0053) (0.0017)   (0.0017)  (0.0052) (0.0016)   (0.0017)
mktrf          -0.1529  1.1153***  -0.0020   -0.2217* 1.0658***  -0.0461
               (0.1257) (0.0403)   (0.0412)  (0.1294) (0.0393)   (0.0409)
hml            0.0965   -0.0881    0.0725    -0.0449  -0.1898*** -0.0180
               (0.2133) (0.0684)   (0.0700)  (0.2232) (0.0678)   (0.0706)
smb            -0.0090  0.6994***  0.1349*   -0.0664  0.6582***  0.0982
               (0.2437) (0.0781)   (0.0800)  (0.2432) (0.0739)   (0.0769)
rmw            -0.2542  -0.4934*** 0.0145    -0.2911  -0.5199*** -0.0091
               (0.2886) (0.0925)   (0.0947)  (0.2865) (0.0870)   (0.0906)
cma            -0.3468  -0.1391    0.0615    -0.1937  -0.0290    0.1595
               (0.3178) (0.1019)   (0.1043)  (0.3244) (0.0985)   (0.1026)
umd                                          -0.3180* -0.2287*** -0.2035***
                                             (0.1615) (0.0491)   (0.0511)
R-squared      0.0232   0.9030     0.0625    0.0474   0.9151     0.1506
R-squared Adj. -0.0085  0.8998     0.0320    0.0100   0.9117     0.1173
N              160      160        160       160      160        160
R2             0.02     0.90       0.06      0.05     0.92       0.15
============================================================================
Standard errors in parentheses.
* p<.1, ** p<.05, ***p<.01
(1): Long-Short Value Weights ~ 5-Factor Fama French Model
(2): Long-Only Rank Weights ~ 5-Factor Fama French Model
(3): Long-Short Rank Weights ~ 5-Factor Fama French Model
(4): Long-Short Value Weights ~ 6-Factor Fama French Model
(5): Long-Only Rank Weights ~ 6-Factor Fama French Model
(6): Long-Short Rank Weights ~ 6-Factor Fama French Model
Annualized Information Ratios:
                       1      2      3      4      5      6
Alpha              0.119  0.010  0.057  0.130  0.018  0.064
Std(resid)         0.215  0.069  0.071  0.213  0.065  0.067
Information Ratio  0.551  0.143  0.801  0.611  0.279  0.949
```

Discussion

Comparison with Strategy 1 -- An Interesting Discovery

Initially, our signal was constructed solely based on the intangibility score. However, after discussions with the professor and reviews of the relevant literature, we decided to integrate a value factor into our signal construction. This modification aims to account for the market's potential anticipation of these risks. Despite this, we also provide results from our original methodology, which omitted current valuations. Notably, this initial approach resulted in a higher alpha for long-short based rank strategies, and it demonstrated more stable returns over a period of time.

Here, we provide two potential explanations. First, determining an accurate valuation factor is particularly challenging in the technology sector, characterized by notably high and volatile multiples. Our current valuation factor construction may not effectively capture these nuances. Future research should aim to develop a more precise valuation factor tailored to technology valuations. Potential approaches could include using a revenue multiple method, which aligns well with the growth-centric nature of tech companies, or implementing a DCF model adjusted for sector-specific risks and growth prospects.

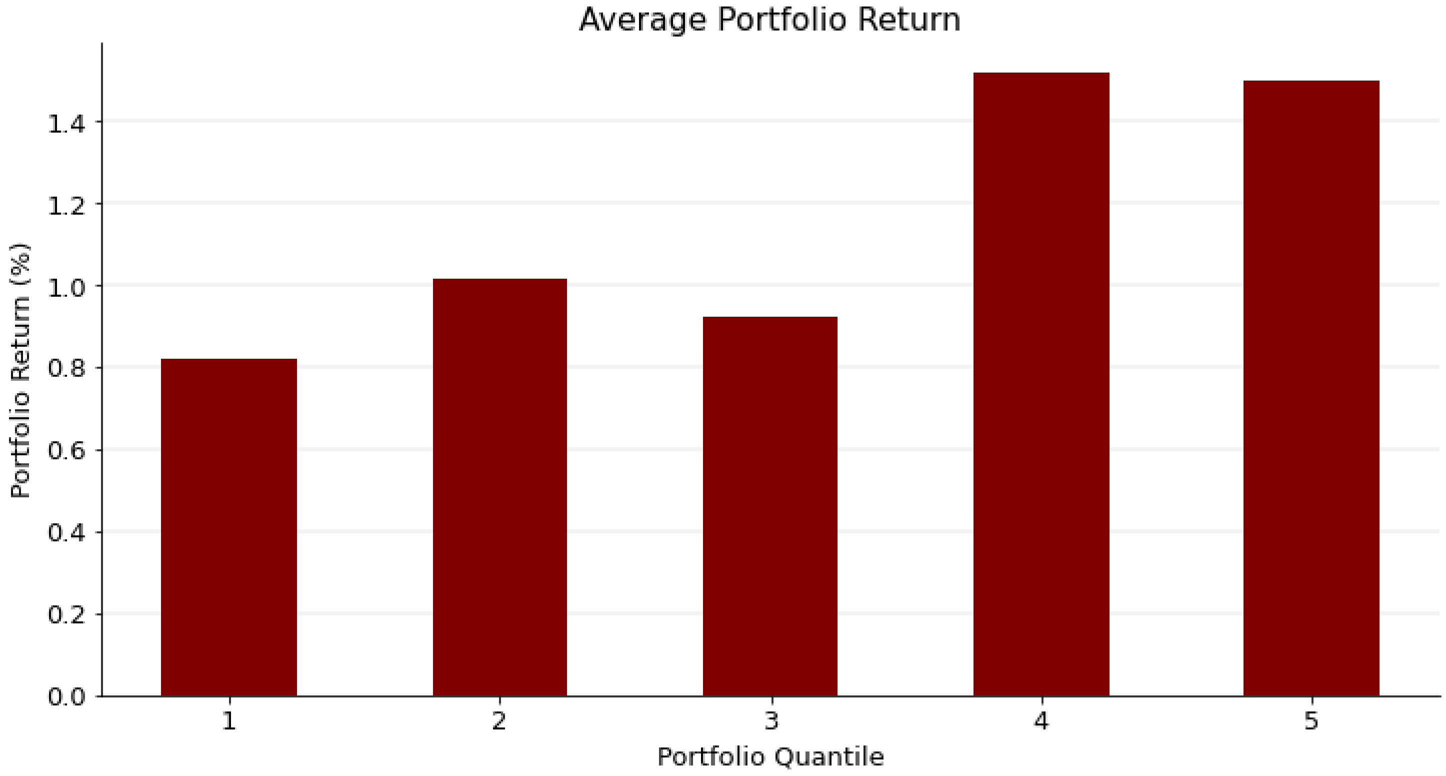
Second, the potential crowdedness of the valuation factor may dilute its returns. In contrast, our findings suggest that the intangibility score factor remains less crowded, possibly because the market often overlooks the impact of intangible assets in the tech industry. For example, there might be over-optimism in M&A deals involving substantial goodwill premiums.

Still, despite the outperformance when excluding the factor, a valuation factor is crucial in our model, because it aligns risk assessment with current stock valuations, allowing for more informed comparisons between inherent risk and market valuation. This ensures our strategy balances risk metrics with a realistic evaluation of each stock's market value.

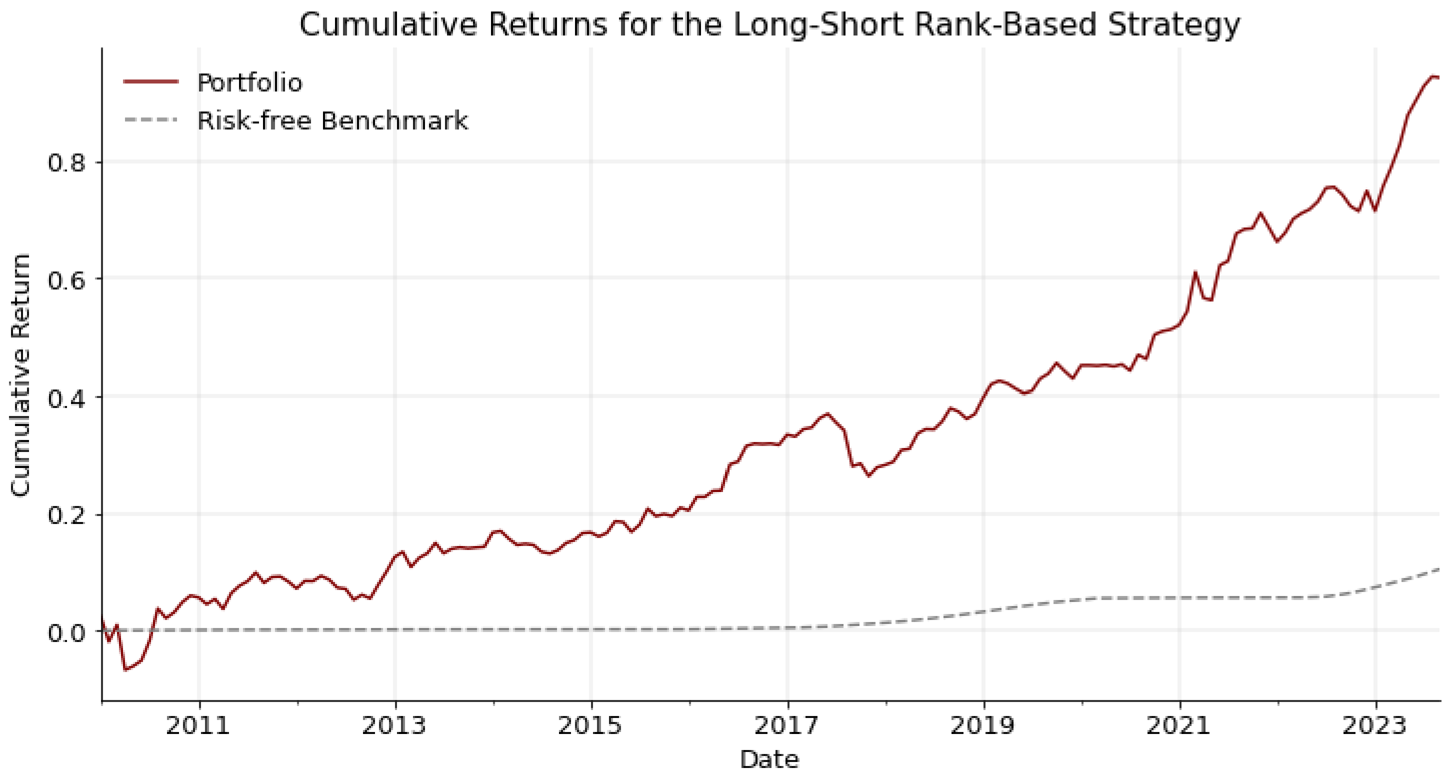
In [14]:

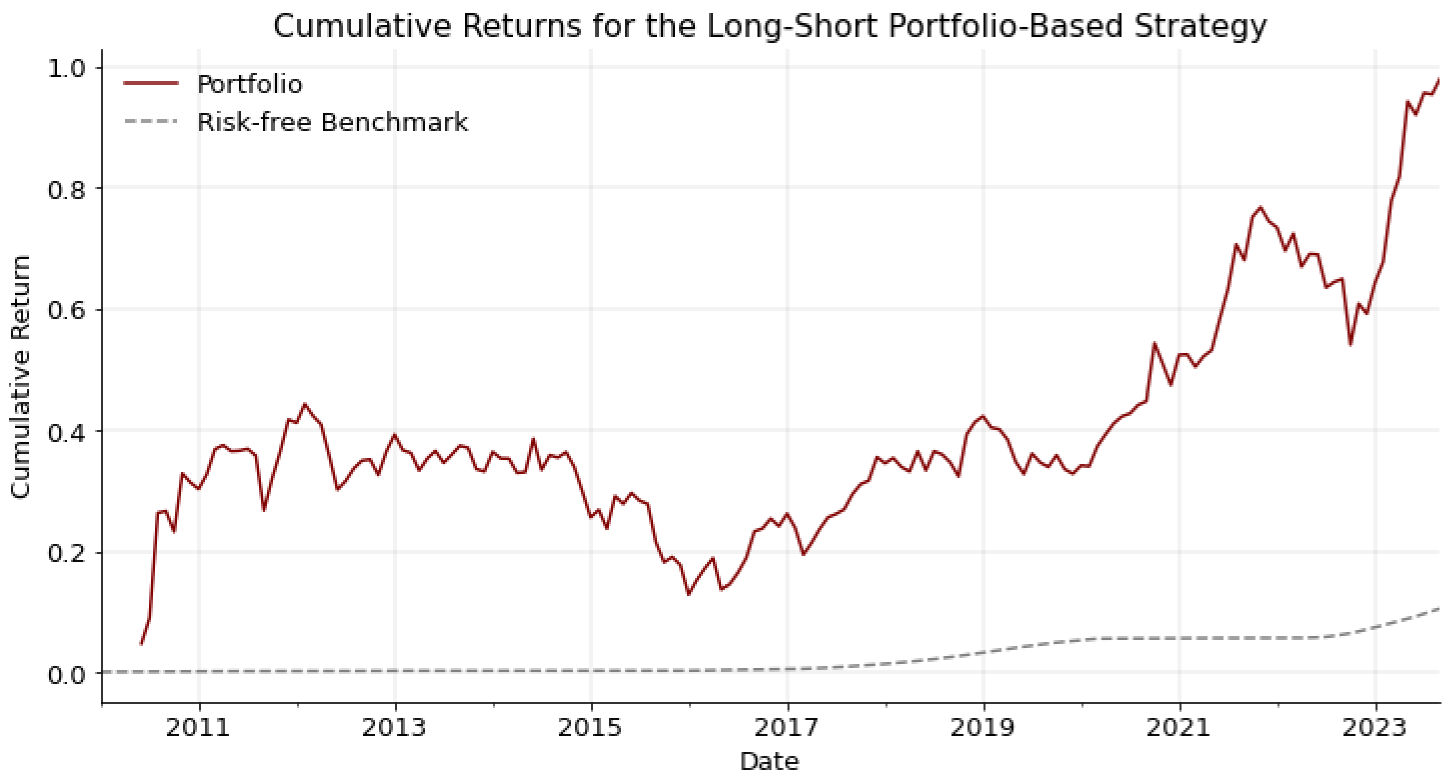
```
# portfolio construction for 'Intan'
_STRATEGY_NAME = 'Intan'
print('> Portfolio construction for' + _STRATEGY_NAME)
if _STRATEGY_NAME == 'Intan':
    df_full['signal'] = df_full['intan_to_noncurrentasset_neg']
elif _STRATEGY_NAME == 'Intan_value':
    df_full['signal'] = df_full['zscore']
else:
    raise Exception('Please provide a valid _STRATEGY_NAME..')
# Select the relevant sample
df = qpm.select_sample(df_full, sample_start = start_date, sample_end = end_date,
                       remove_micro_caps = _REMOVE_MICRO_CAPS)
print('> Select Industry for Analysis...')
df_industry = df_full[df_full[_INDUSTRY_REFERENCE]==_INDUSTRY_CHOSEN]

df_industry1, df_rets1 = qpm_modified.create_portfolios_modified(df_industry, sort_frequency = _SORT_FREQUENCY, num_po

qpm.analyze_strategy(df_rets1, analysis_type = 'Performance')

# Include fama-french 5 factor industry performance for comparsion
ff_factor = 'F-F_Research_Data_5_Factors_2x3'
ff_factor_data = web.DataReader(ff_factor, 'famafrench', start=start_date, end=end_date)[0]
ff_portfolio = '49_Industry_Portfolios'
ff_portfolio_data = web.DataReader(ff_portfolio, 'famafrench', start=start_date, end=end_date)[0]
ff_portfolio_data = ff_portfolio_data.sub(ff_factor_data.RF, axis=0)/100
sectors = ['Hardw', 'Softw', 'Chips', 'LabEq', 'ElcEq', 'Telcm']
ff_portfolio_data['ff Sector Performance'] = ff_portfolio_data[sectors].mean(axis=1)
ff_portfolio_data = ff_portfolio_data.reset_index().sort_values(['Date']).rename(columns = {'Date' : 'ldate'})
ff_portfolio_data['ldate'] = ff_portfolio_data['ldate'].dt.strftime('%Y-%m') + '-01'
ff_portfolio_data['ldate'] = pd.to_datetime(ff_portfolio_data['ldate'])

df2 = pd.merge(ff_portfolio_data, df_rets1, on = ['ldate'], validate = 'many_to_one', indicator = True).reset_index()
df2.sort_values(['ldate'], inplace = True)
df2['CLNretP'] = np.log(1 + df2['retP_rank_longonly']).cumsum()
plot_df = df2[[ 'CLNretP', 'ff Sector Performance','ldate']].set_index('ldate')
fig = plt.figure(figsize = (12, 6))
plot_df['CLNretP'].plot(ax = plt.gca(), color = 'maroon', label = 'Portfolio')
plot_df['ff Sector Performance'].cumsum().plot(ax = plt.gca(), color = 'gray', label = 'Industry Fama French')
plt.xlabel('Date'); plt.ylabel('Cumulative Return')
plt.title('Cumulative Returns for the Long-Only Strategy')
plt.legend()
plt.show()
plt.close()

qpm.analyze_strategy(df_rets1, analysis_type = 'Summary')

qpm.analyze_strategy(df_rets1, analysis_type = 'Factor Regression')
```

```
> Portfolio construction forIntan
> Selecting Sample for Given Criteria...
> Select Industry for Analysis...
> Sorting stocks into 5 portfolios at frequency: Quarterly...
> Computing returns using various weights...
> Merging strategy returns with Fama and French factor returns...
```

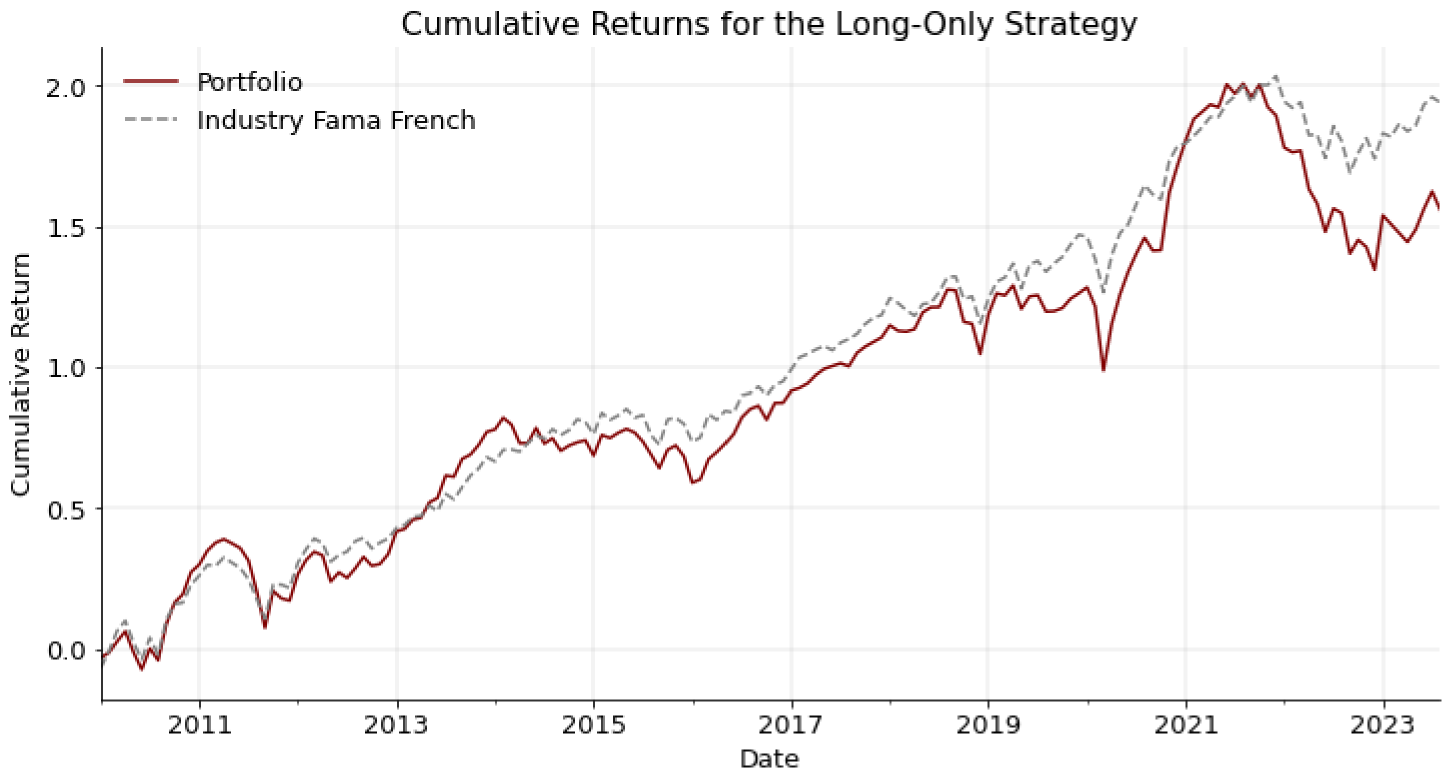

Average Portfolio Return
Portfolio Return (%)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1
2
3
4
5
Portfolio Quantile


Cumulative Returns for the Long-Only Strategy
Portfolio
Market
Cumulative Return
0.0
0.5
1.0
1.5
2.0
2011
2013
2015
2017
2019
2021
2023
Date


Cumulative Returns for the Long-Short Rank-Based Strategy
Portfolio
Risk-free Benchmark
Cumulative Return
0.0
0.2
0.4
0.6
0.8
2011
2013
2015
2017
2019
2021
2023
Date



Cumulative Returns for the Long-Short Portfolio-Based Strategy
Portfolio
Risk-free Benchmark
Cumulative Return
Date



Cumulative Returns for the Long-Only Strategy
Portfolio
Industry Fama French
Cumulative Return
Date

```
> Merging strategy returns with Fama and French factor returns...
                     count      mean       std       min       25%       50%  \
retP_rank_longonly   165.0  0.010854  0.062577 -0.203170 -0.025842  0.013910
retP_rank_longshort  165.0  0.005286  0.019732 -0.074254 -0.004844  0.004648
retF_vw              160.0  0.006196  0.037843 -0.105665 -0.016999  0.005059
mktrf                165.0  0.010054  0.044228 -0.133900 -0.015500  0.012900
smb                  165.0 -0.000310  0.025981 -0.082800 -0.018200  0.001200
hml                  165.0 -0.001476  0.033048 -0.138700 -0.018600 -0.004200

                          75%       max
retP_rank_longonly   0.045310  0.225628
retP_rank_longshort  0.015739  0.069712
retF_vw              0.023447  0.189519
mktrf                0.034400  0.136500
smb                  0.015700  0.071100
hml                  0.013700  0.127500
> Merging strategy returns with Fama and French factor returns...

---------------------------------------------------------------
> Running Factor Regressions: Table 1  - 3 Fama-French Factors
---------------------------------------------------------------

==========================================================================
                 (1)      (2)       (3)       (4)       (5)        (6)
--------------------------------------------------------------------------
const          0.0057*  -0.0026   0.0061*** 0.0055*   -0.0004    0.0060***
               (0.0031) (0.0024)  (0.0015)  (0.0031)  (0.0017)   (0.0015)
mktrf          0.0435   1.2713*** -0.0271   0.0368    1.0904***  -0.0359
               (0.0684) (0.0531)  (0.0334)  (0.0727)  (0.0403)   (0.0356)
hml                                         -0.2217** -0.2623*** -0.0942**
                                            (0.0928)  (0.0514)   (0.0454)
smb                                         0.0745    0.8870***  0.0589
                                            (0.1273)  (0.0706)   (0.0623)
R-squared      0.0025   0.7839    0.0042    0.0378    0.8938     0.0319
R-squared Adj. -0.0038  0.7825    -0.0021   0.0193    0.8917     0.0133
N              160      160       160       160       160        160
R2             0.00     0.78      0.00      0.04      0.89       0.03
==========================================================================
Standard errors in parentheses.
* p<.1, ** p<.05, ***p<.01
(1): Long-Short Value Weights ~ CAPM Model
(2): Long-Only Rank Weights ~ CAPM Model
(3): Long-Short Rank Weights ~ CAPM Model
(4): Long-Short Value Weights ~ 3-Factor Fama French Model
(5): Long-Only Rank Weights ~ 3-Factor Fama French Model
(6): Long-Short Rank Weights ~ 3-Factor Fama French Model
Annualized Information Ratios:
                       1      2      3      4      5      6
Alpha              0.069 -0.031  0.073  0.065 -0.005  0.073
Std(resid)         0.131  0.101  0.064  0.128  0.071  0.063
Information Ratio  0.528 -0.303  1.147  0.511 -0.072  1.156



---------------------------------------------------------------
> Running Factor Regressions: Table 2  - 5 Fama-French Factors + Momentum
---------------------------------------------------------------

==========================================================================
                 (1)      (2)        (3)      (4)       (5)        (6)
--------------------------------------------------------------------------
const          0.0065** 0.0011     0.0059*** 0.0068** 0.0017     0.0060***
               (0.0031) (0.0016)   (0.0015)  (0.0031) (0.0015)   (0.0015)
mktrf          0.0301   1.1051***  -0.0429   0.0028   1.0653***  -0.0481
               (0.0741) (0.0376)   (0.0365)  (0.0768) (0.0373)   (0.0380)
hml            -0.0823  -0.1288**  -0.0903   -0.1384  -0.2105*** -0.1009
               (0.1257) (0.0638)   (0.0619)  (0.1324) (0.0644)   (0.0656)
smb            -0.0258  0.6881***  0.0897    -0.0485  0.6550***  0.0854
               (0.1436) (0.0729)   (0.0708)  (0.1443) (0.0702)   (0.0715)
rmw            -0.1950  -0.4749*** 0.0886    -0.2096  -0.4962*** 0.0858
               (0.1700) (0.0863)   (0.0838)  (0.1700) (0.0826)   (0.0842)
cma            -0.2579  -0.1631*   -0.0346   -0.1972  -0.0746    -0.0231
               (0.1873) (0.0950)   (0.0923)  (0.1924) (0.0936)   (0.0953)
umd                                          -0.1262  -0.1838*** -0.0239
                                             (0.0958) (0.0466)   (0.0475)
R-squared      0.0580   0.9129     0.0396    0.0686   0.9209     0.0412
R-squared Adj. 0.0274   0.9101     0.0084    0.0321   0.9178     0.0036
N              160      160        160       160      160        160
R2             0.06     0.91       0.04      0.07     0.92       0.04
==========================================================================
Standard errors in parentheses.
* p<.1, ** p<.05, ***p<.01
(1): Long-Short Value Weights ~ 5-Factor Fama French Model
(2): Long-Only Rank Weights ~ 5-Factor Fama French Model
(3): Long-Short Rank Weights ~ 5-Factor Fama French Model
(4): Long-Short Value Weights ~ 6-Factor Fama French Model
(5): Long-Only Rank Weights ~ 6-Factor Fama French Model
(6): Long-Short Rank Weights ~ 6-Factor Fama French Model
Annualized Information Ratios:
                       1      2      3      4      5      6
Alpha              0.078  0.013  0.071  0.082  0.020  0.071
Std(resid)         0.127  0.064  0.063  0.126  0.061  0.062
Information Ratio  0.611  0.208  1.129  0.650  0.325  1.144
```