BEM114_HW3_Andrew_Daniel_Kyle

April 26, 2024

1 BEM114 Homework 3 - Value of Intangibles

Names: Andrew Zabelo, Daniel Wen, Kyle McCandless
Student IDs: 2176083, 2159859, 2157818

1.1 Setup

1.1.1 Imports and Helper Functions

[1]:
```
import pandas as pd
import numpy as np

from datetime import datetime
from dateutil.relativedelta import relativedelta

import statsmodels.api as sm
import matplotlib.pyplot as plt
```

[2]:
```
# Gets the next month from a date string
def get_next_month_str(d1):
    return (datetime.strptime(d1, "%Y-%m") + relativedelta(months=1)).
 ↪strftime("%Y-%m")

def is_next_month(d1, d2):
    return d2 == get_next_month_str(d1)

# Briefly test is_next_month
assert is_next_month('1986-05', '1986-06')
assert is_next_month('2012-12', '2013-01')
assert is_next_month('1999-12', '2000-01')
assert not is_next_month('1986-04', '1986-06')
assert not is_next_month('1998-04', '1999-05')
assert not is_next_month('1984-01', '1986-10')

# Given a group of stocks, calculate equal-weighted and value-weighted weights
def calc_weights(group):
    if group['rebalance'].sum() > 0:
        # Calc equal weights
```

```
        group['weights_eq'] = 1 / float(group['PERMNO'].count())
        assert(group['PERMNO'].count() == group['PERMNO'].nunique())
        # Calc value weights
        group['weights_val'] = group['MV'] / group['MV'].sum()
        return group
    else:
        group['weights_eq'] = np.nan
        group['weights_val'] = np.nan
        return group

# Adds a final month of 0 returns any time a stock becomes de-listed
def add_padding_month(group):
    # Sorted list of dates in this group
    dates_list = sorted([''.join(str_list) for str_list in group['date'].
 ↪agg(list)])

    # Collect de-listing dates
    dates_to_add = [group['date'].max()]
    prev_date = dates_list[0]
    for date in dates_list[1:]:
        if not is_next_month(prev_date, date):
            dates_to_add.append(prev_date)
        prev_date = date

    # Add new padding rows for the month after each de-listing date
    new_rows = []
    for date in dates_to_add:
        new_row = pd.Series({'company': group['company'].iloc[0],
                             'PERMNO': group['PERMNO'].iloc[0],
                             'date': get_next_month_str(date),
                             'RET': 0.0,
                             'MV': 1.0})
        new_rows.append(new_row)

    return pd.concat([group, pd.DataFrame(new_rows)], ignore_index=True)

# Calculates returns and prints the returns mean, vol, and Sharpe ratio for a␣
 ↪strategy
def analyze(returns, strat_name):
    strat_mean = returns.mean()
    strat_vol = returns.std()
    strat_sharpe = strat_mean / strat_vol
    print(f"{strat_name} monthly returns:\nMean = {strat_mean}%\nVolatility =␣
 ↪{strat_vol}%\nSharpe Ratio = {strat_sharpe}")

# Estimates the CAPM and FF3 models on df_old using the returns found in␣
 ↪ret_col_name
```

```python
def estimate_models(df_old, return_col_name, ff5_mom):
    df = pd.merge(df_old, ff5_mom, how='inner', on=['date'])
    assert(len(df) == len(df_old))

    # Estimate CAPM
    print('CAPM')
    capm_model = sm.OLS(df[return_col_name] - df['RF'], sm.add_constant(df[['Mkt-RF']])).fit()
    capm_beta = capm_model.params['Mkt-RF']
    print(capm_model.summary())

    # Estimate FF3
    print('FF3')
    print(sm.OLS(df[return_col_name] - df['RF'], sm.add_constant(df[['Mkt-RF', 'SMB', 'HML']])).fit().summary())

    # Estimate Carhart
    print('Carhart')
    print(sm.OLS(df[return_col_name] - df['RF'], sm.add_constant(df[['Mkt-RF', 'SMB', 'HML', 'MOM']])).fit().summary())

    # Estimate FF5
    print('FF5')
    print(sm.OLS(df[return_col_name] - df['RF'], sm.add_constant(df[['Mkt-RF', 'SMB', 'HML', 'RMW', 'CMA']])).fit().summary())

    return capm_beta

# Plots the cumulative returns for a strategy versus the CAPM-implied returns
def plot_cum_returns(df, return_col_name, capm_beta, title):
    df[return_col_name + '_MIR'] = df['RF'] + capm_beta * df['Mkt-RF']

    dates = df['date'] // 100 + (df['date'] % 100) / 12
    strategy_cumulative = (df[return_col_name] / 100 + 1.0).cumprod()
    mir_cumulative = (df[return_col_name + '_MIR'] / 100 + 1.0).cumprod()
    market_cumulative = ((df['Mkt-RF'] + df['RF']) / 100 + 1.0).cumprod()

    plt.figure()
    plt.plot(dates, strategy_cumulative, label=f'{return_col_name} Portfolio Value')
    plt.plot(dates, mir_cumulative, label=f'CAPM-Implied Portfolio Value')
    plt.plot(dates, market_cumulative, label=f'Market Portfolio Value')

    plt.title(f'{title} Model Performance')
    plt.xlabel('Date')
    plt.ylabel('Cumulative Portfolio Value')
```

```
    plt.legend()
    plt.show()
```

1.1.2 Process Dataframes

[3]:
```
'''
Load CRSP data
'''

crsp = pd.read_csv('crsp_1926_2020.zip')

# Convert prices and returns to numeric and drop NaNs
crsp['PRC'] = pd.to_numeric(crsp['PRC'], errors='coerce')
crsp['RET'] = pd.to_numeric(crsp['RET'], errors='coerce')
crsp = crsp.dropna(subset=['PRC', 'RET'])

# Set types for relevant columns
crsp = crsp.astype({'date': 'string', 'SHRCD': 'int', 'EXCHCD': 'int'})

# [From HW2] Filter SHRCD and EXCHCD
crsp = crsp[crsp['SHRCD'].isin([10, 11])]
crsp = crsp[crsp['EXCHCD'].isin([1, 2, 3])]

# Reformat date column and add market value column
crsp['date'] = crsp['date'].str[:-3]
crsp['year'] = crsp['date'].str[:-3].astype('int')
crsp['MV'] = np.abs(crsp['PRC']) * crsp['SHRCD']
crsp['RET'] *= 100
crsp
```

[3]:
```
         PERMNO     date  SHRCD  EXCHCD        PRC      RET    SHROUT  year  \
2         10000  1986-02     10       3   -3.25000 -25.7143    3680.0  1986
3         10000  1986-03     10       3   -4.43750  36.5385    3680.0  1986
4         10000  1986-04     10       3   -4.00000  -9.8592    3793.0  1986
5         10000  1986-05     10       3   -3.10938 -22.2656    3793.0  1986
6         10000  1986-06     10       3   -3.09375  -0.5025    3793.0  1986
…           …        …    …      …          …        …    …
4705164   93436  2020-08     11       3  498.32001  74.1452  931809.0  2020
4705165   93436  2020-09     11       3  429.01001 -13.9087  948000.0  2020
4705166   93436  2020-10     11       3  388.04001  -9.5499  947901.0  2020
4705167   93436  2020-11     11       3  567.59998  46.2736  947901.0  2020
4705168   93436  2020-12     11       3  705.66998  24.3252  959854.0  2020

                   MV
2        1.196000e+04
3        1.633000e+04
4        1.517200e+04
```

```
5          1.179388e+04
6          1.173459e+04
…                    …
4705164   4.643391e+08
4705165   4.067015e+08
4705166   3.678235e+08
4705167   5.380286e+08
4705168   6.773402e+08

[3563041 rows x 9 columns]
```

[4]:

```python
'''
Load FF5 and Industries data
'''

ff5 = pd.read_csv('ff5_factors.csv')
ff5 = ff5.astype({'date': 'string'})
ff5['date'] = ff5['date'].apply(lambda x: x[:4] + '-' + x[4:])

mom = pd.read_csv('F-F_Momentum_Factor.CSV')
mom = mom.astype({'date': 'string'})
mom['date'] = mom['date'].apply(lambda x: x[:4] + '-' + x[4:])

ff5_mom = pd.merge(ff5, mom, on='date', how='inner')
assert len(ff5_mom) == len(ff5)
ff5_mom
```

[4]:

```
        date  Mkt-RF   SMB   HML   RMW   CMA    RF   MOM
0    1963-07   -0.39 -0.41 -0.97  0.68 -1.18  0.27  0.90
1    1963-08    5.07 -0.80  1.80  0.36 -0.35  0.25  1.01
2    1963-09   -1.57 -0.52  0.13 -0.71  0.29  0.27  0.19
3    1963-10    2.53 -1.39 -0.10  2.80 -2.01  0.29  3.12
4    1963-11   -0.85 -0.88  1.75 -0.51  2.24  0.27 -0.74
..       …      …     …     …     …     …     …     …
723  2023-10   -3.19 -4.05  0.19  2.46 -0.66  0.47  1.73
724  2023-11    8.84 -0.12  1.64 -3.91 -1.00  0.44  2.75
725  2023-12    4.87  7.32  4.93 -3.07  1.32  0.43 -5.51
726  2024-01    0.71 -5.74 -2.38  0.69 -0.96  0.47  5.18
727  2024-02    5.06 -0.78 -3.48 -1.98 -2.13  0.42  4.92

[728 rows x 8 columns]
```

[5]:

```python
'''
Load 100 Best Companies to Work for in America
'''

bcw = pd.read_csv('bcwlist_modified.csv')
```

```python
bcw = bcw.dropna(subset=['permno'])

# Set types for relevant columns
bcw = bcw.astype({'rank': 'int', 'company': 'string', 'year': 'int'})
bcw.rename(columns={'permno': 'PERMNO'}, inplace=True)
bcw = bcw.sort_values(by=['year', 'rank'])
bcw
```

[5]:

```
      rank                          company   PERMNO  year
0        1          AT&T Bell Laboratories  66093.0  1984
1        2           Trammell Crow Company  85629.0  1984
2        3                  Delta Airlines  26112.0  1984
3        4                 Federal Express  60628.0  1984
4        5                   Goldman Sachs  86868.0  1984
…     …                                 …        …     …
2486    87                          AbbVie  13721.0  2020
2487    88  Encompass Home Health & Hospice  10693.0  2020
2493    94                   Goldman Sachs  86868.0  2020
2498    99                  Delta Airlines  91926.0  2020
2499   100             Four Seasons Hotels  84592.0  2020

[1329 rows x 4 columns]
```

1.2 Problem 1

1.2.1 Part A - Process

Data cleaning: * Filter CRSP to ordinary/common shares (SHRCD = 10 or 11) and NYSE, AMEX, and NASDAQ stocks (EXCHCD = 1, 2, or 3)] * Do not filter out CRSP negative prices. We only need the estimates of return data. * Assume that if no BCW PERMNO then not publicly traded and remove from dataset * Augment bcw dataframe with all years a portfolio will be active. For example, the portfolio formed in 1984 is active in 1985, 1986, … 1992.

Calculating weights: * Merge CRSP data with bcw dataframe on year, so that monthly price data for all stocks in an active best companies portfolio are present * Create **rebalance column:** True if rebalancing based on stock prices that month, false otherwise. Rebalancing is true iff it is January the year the portfolio was formed or a stock is listed or delisted in that month. * Apply calc_weights, which rebalances the entire portfolio if rebalancing == True, and if not returns NaN weights. * After sorting by PERMNO and date Fill NaN weights with the first non-NaN weights above them, keeping the weights the same for months where we don't rebalance the portfolio. One stock's weights can never be filled with the weights of another since the earliest date of any stock always has rebalance == True * Pad all stocks that are de-listed at any time(s) with an extra month of 0 returns * Shift weights down by one date for each stock. For de-listing months, the previous month's weights occupy the padded month as desired, since the portfolio will carry the stock with 0 returns during the month it is de-listed, then be updated next month. For listing months, the weights from the first month the stock is present are shifted down to form the portfolio for the next month

Calculating returns: * As in HW2, multiply the returns of each stock with its lagged weights, which are the weights the portfolio started that month with

[6]:
```python
# Prepare bcw for merge by adding years between 1984 - 1993, 1993 - 1998
bcw['year_formed'] = bcw['year']

bcw_extra = []
for year_formed, gap_length in zip([1984, 1993], [1993 - 1984, 1998 - 1993]):
    to_increment = bcw[bcw['year_formed'] == year_formed]
    for increment in range(1, gap_length):
        for _, row in to_increment.iterrows():
            row['year'] += increment
            bcw_extra.append(row.to_dict())

bcw_extra = pd.DataFrame(bcw_extra, columns=bcw.columns)
bcw = pd.concat([bcw, bcw_extra], ignore_index=True)
bcw
```

[6]:

	rank	company	PERMNO	year	year_formed
0	1	AT&T Bell Laboratories	66093.0	1984	1984
1	2	Trammell Crow Company	85629.0	1984	1984
2	3	Delta Airlines	26112.0	1984	1984
3	4	Federal Express	60628.0	1984	1984
4	5	Goldman Sachs	86868.0	1984	1984
…	…	…	…	…	…
2320	101	Viking Freight System	80814.0	1997	1993
2321	101	Wal-Mart Stores	55976.0	1997	1993
2322	101	Weyerhaeuser Company	39917.0	1997	1993
2323	101	Worthington Industries	83601.0	1997	1993
2324	101	Xerox	27983.0	1997	1993

[2325 rows x 5 columns]

[7]:
```python
# Merge bcw and crsp
df = pd.merge(bcw, crsp, how='inner', on=['year', 'PERMNO'])
df = df.sort_values(by=['PERMNO', 'date'])

# Find dates where firms were listed and de-listed
special_dates = set(df.groupby('PERMNO')['date'].min().tolist())
last_trade_dates = df.groupby('PERMNO')['date'].max().tolist()
for date in last_trade_dates:
    special_dates.add(get_next_month_str(date))

# Add to special dates if any firms are de-listed then listed again
prev_row = df.iloc[0]
for _, row in df.iloc[1:].iterrows():
    if (not is_next_month(prev_row['date'], row['date'])) and␣
↪(prev_row['PERMNO'] == row['PERMNO']):
```

```
        special_dates.add(get_next_month_str(prev_row['date']))
        special_dates.add(row['date'])
    prev_row = row

# Add rebalance column -- rebalance if its January when the portfolio was␣
 ↪formed or if its a special date
df['rebalance'] = (df['date'].str[-2:] == '01') | (df['date'].
 ↪isin(special_dates))

df = df.drop(['rank', 'SHRCD', 'EXCHCD', 'PRC', 'SHROUT'], axis=1)
df = df.sort_values(by=['date'])
df
```

[7]:

```
                                    company   PERMNO  year  year_formed  \
562                                    Moog  61807.0  1984         1984
394                    Inland Steel Company  12458.0  1984         1984
490                     Liebert Corporation  49411.0  1984         1984
154                               Armstrong  19692.0  1984         1984
70                                     Time  40483.0  1984         1984
…                                         …        …   …            …
13271                                Hilton  14338.0  2020         2020
13379                                CarMax  89508.0  2020         2020
13331                      American Express  59176.0  2020         2020
13391                 Capital One Financial  81055.0  2020         2020
13583  First American Financial Corporation  93374.0  2020         2020

          date      RET            MV  rebalance
562    1984-01 -15.0000  8.582875e+04       True
394    1984-01  -3.2258  7.469400e+05       True
490    1984-01   1.7857  3.116048e+05       True
154    1984-01  -5.4299  6.469856e+05       True
70     1984-01  -8.9417  2.671005e+06       True
…          …       …            …         …
13271  2020-12   7.3627  3.086864e+07      False
13379  2020-12   1.0483  1.540019e+07      False
13331  2020-12   1.9563  9.735697e+07      False
13391  2020-12  15.4250  4.521369e+07      False
13583  2020-12   7.5351  5.764231e+06      False

[23034 rows x 8 columns]
```

[8]:

```
# Group by date and calculate weights
df_weights = df.groupby('date', group_keys=False).apply(calc_weights)

# Assert that calc_weights is returning weights when rebalance is needed only
assert len(df_weights[(df_weights['rebalance'] == True) & (np.
 ↪isnan(df_weights['weights_eq']))]) == 0
```

```
assert len(df_weights[(df_weights['rebalance'] == True) & (np.
 ↪isnan(df_weights['weights_val']))]) == 0

# Fill the NaNs returned from calc_weights when there are no rebalances
# using the weights for that PERMNO on the previous date
df_weights = df_weights.sort_values(['PERMNO', 'date'])
df_weights['weights_eq'] = df_weights['weights_eq'].fillna(method='ffill')
df_weights['weights_val'] = df_weights['weights_val'].fillna(method='ffill')

# Pad all stocks with 1 extra month of 0 returns for weight shift
df_weights = df_weights.groupby('PERMNO').apply(add_padding_month).
 ↪reset_index(drop=True)

# Shift weights
df_weights = df_weights.sort_values(['PERMNO', 'date'])
df_weights['weights_eq_lag'] = df_weights.groupby('PERMNO')['weights_eq'].
 ↪shift(1)
df_weights['weights_val_lag'] = df_weights.groupby('PERMNO')['weights_val'].
 ↪shift(1)
df_weights = df_weights.dropna(subset=['weights_eq_lag', 'weights_val_lag'])

# Assert that weights add up to one for all dates
test1 = df_weights.groupby('date')['weights_eq_lag'].sum()
test2 = df_weights.groupby('date')['weights_val_lag'].sum()
assert test1.apply(lambda x: np.isclose(x, 1.0, atol=0.00001)).all()
assert test2.apply(lambda x: np.isclose(x, 1.0, atol=0.00001)).all()

df_weights = df_weights.drop(['year', 'MV'], axis=1)
df_weights = df_weights.sort_values(['date', 'PERMNO'])
df_weights
```

[8]:

```
                                     company   PERMNO  year_formed     date  \
402            Atlantic Richfield Company  10604.0       1984.0  1984-02
710                      Dana Corporation  11607.0       1984.0  1984-02
893                               Du Pont  11703.0       1984.0  1984-02
1062               Eastman Kodak Company  11754.0       1984.0  1984-02
1171                   Exxon Corporation  11850.0       1984.0  1984-02
…                                       …        …            …        …
22660                     Salesforce.com  90215.0          NaN  2021-01
23068                     Delta Airlines  91926.0          NaN  2021-01
23105                        T-Mobile US  91937.0          NaN  2021-01
23305          Hyatt Hotels Corporation  93098.0          NaN  2021-01
23366  First American Financial Corporation  93374.0          NaN  2021-01

           RET  rebalance  weights_eq  weights_val  weights_eq_lag  \
402    -0.5525      False    0.014085     0.038332        0.014085
710   -13.3739      False    0.014085     0.005508        0.014085
```

```
893     -3.3668      False    0.014085    0.040308        0.014085
1062    -5.8319      False    0.014085    0.040958        0.014085
1171    -1.1321      False    0.014085    0.114076        0.014085
…         …         …         …          …             …
22660    0.0000       <NA>         NaN         NaN        0.029412
23068    0.0000       <NA>         NaN         NaN        0.029412
23105    0.0000       <NA>         NaN         NaN        0.029412
23305    0.0000       <NA>         NaN         NaN        0.029412
23366    0.0000       <NA>         NaN         NaN        0.029412

       weights_val_lag
402           0.038332
710           0.005508
893           0.040308
1062          0.040958
1171          0.114076
…                  …
22660         0.085811
23068         0.019130
23105         0.036010
23305         0.001641
23366         0.003695

[23034 rows x 10 columns]
```

1.3 Problem 2

1.3.1 Part A

[9]:

```
# Add returns
df_weights['weighted_eq_ret'] = df_weights['weights_eq_lag'] * df_weights['RET']
df_weights['weighted_val_ret'] = df_weights['weights_val_lag'] *␣
↪df_weights['RET']

eq_returns = df_weights.groupby('date')['weighted_eq_ret'].sum()
val_returns = df_weights.groupby('date')['weighted_val_ret'].sum()

analyze(eq_returns, "Equal-weighted Best Companies")
print()
analyze(val_returns, "Value-weighted Best Companies")
```

```
Equal-weighted Best Companies monthly returns:
Mean = 1.220920902080068%
Volatility = 5.403471378171253%
Sharpe Ratio = 0.22595121110704866

Value-weighted Best Companies monthly returns:
Mean = 1.0771412191404004%
```

```
Volatility = 5.264385220495521%
Sharpe Ratio = 0.20460911844878485
```

1.3.2 Part B

[10]:
```python
eq_capm_beta = estimate_models(eq_returns, 'weighted_eq_ret', ff5_mom)
```

```
CAPM
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.892
Model:                            OLS   Adj. R-squared:                  0.892
Method:                 Least Squares   F-statistic:                     3653.
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          8.79e-216
Time:                        10:59:49   Log-Likelihood:                -885.39
No. Observations:                 444   AIC:                             1775.
Df Residuals:                     442   BIC:                             1783.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.1085      0.086      1.267      0.206      -0.060       0.277
Mkt-RF         1.1413      0.019     60.437      0.000       1.104       1.178
==============================================================================
Omnibus:                       58.357   Durbin-Watson:                   1.962
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              168.101
Skew:                           0.615   Prob(JB):                     3.14e-37
Kurtosis:                       5.752   Cond. No.                         4.60
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
FF3
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.901
Model:                            OLS   Adj. R-squared:                  0.900
Method:                 Least Squares   F-statistic:                     1334.
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          1.93e-220
Time:                        10:59:49   Log-Likelihood:                -866.28
No. Observations:                 444   AIC:                             1741.
Df Residuals:                     440   BIC:                             1757.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
```

```
------------------------------------------------------------------------------
const          0.1105      0.082      1.340      0.181      -0.052       0.272
Mkt-RF         1.1183      0.019     59.270      0.000       1.081       1.155
SMB            0.1755      0.028      6.205      0.000       0.120       0.231
HML            0.0325      0.028      1.167      0.244      -0.022       0.087
==============================================================================
Omnibus:                       66.589   Durbin-Watson:                   2.034
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              168.211
Skew:                           0.749   Prob(JB):                     2.97e-37
Kurtosis:                       5.617   Cond. No.                         4.74
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
Carhart
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.911
Model:                            OLS   Adj. R-squared:                  0.911
Method:                 Least Squares   F-statistic:                     1128.
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          2.21e-229
Time:                        10:59:49   Log-Likelihood:                -841.73
No. Observations:                 444   AIC:                             1693.
Df Residuals:                     439   BIC:                             1714.
Df Model:                           4
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.2103      0.079      2.652      0.008       0.054       0.366
Mkt-RF         1.0836      0.019     58.521      0.000       1.047       1.120
SMB            0.1743      0.027      6.504      0.000       0.122       0.227
HML           -0.0177      0.027     -0.649      0.517      -0.071       0.036
MOM           -0.1299      0.018     -7.165      0.000      -0.165      -0.094
==============================================================================
Omnibus:                       30.015   Durbin-Watson:                   2.080
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               59.713
Skew:                           0.397   Prob(JB):                     1.08e-13
Kurtosis:                       4.612   Cond. No.                         5.19
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
FF5
                            OLS Regression Results
==============================================================================
```

```
Dep. Variable:                      y   R-squared:                       0.905
Model:                            OLS   Adj. R-squared:                  0.904
Method:                 Least Squares   F-statistic:                     836.2
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          1.87e-221
Time:                        10:59:49   Log-Likelihood:                -856.62
No. Observations:                 444   AIC:                             1725.
Df Residuals:                     438   BIC:                             1750.
Df Model:                           5
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.1057      0.084      1.254      0.210      -0.060       0.271
Mkt-RF         1.1089      0.020     54.639      0.000       1.069       1.149
SMB            0.2160      0.030      7.121      0.000       0.156       0.276
HML            0.0760      0.037      2.056      0.040       0.003       0.149
RMW            0.1123      0.040      2.831      0.005       0.034       0.190
CMA           -0.1696      0.057     -2.998      0.003      -0.281      -0.058
==============================================================================
Omnibus:                       53.301   Durbin-Watson:                   2.037
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              103.410
Skew:                           0.692   Prob(JB):                     3.51e-23
Kurtosis:                       4.917   Cond. No.                         5.30
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

[11]:
```python
val_capm_beta = estimate_models(val_returns, 'weighted_val_ret', ff5_mom)
```

```
CAPM
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.834
Model:                            OLS   Adj. R-squared:                  0.834
Method:                 Least Squares   F-statistic:                     2220.
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          1.90e-174
Time:                        10:59:49   Log-Likelihood:                -968.65
No. Observations:                 444   AIC:                             1941.
Df Residuals:                     442   BIC:                             1949.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0145      0.103      0.140      0.888      -0.189       0.218
Mkt-RF         1.0734      0.023     47.119      0.000       1.029       1.118
```

```
==============================================================================
Omnibus:                       27.575   Durbin-Watson:                   2.132
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               95.506
Skew:                           0.008   Prob(JB):                     1.82e-21
Kurtosis:                       5.272   Cond. No.                         4.60
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
FF3
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.854
Model:                            OLS   Adj. R-squared:                  0.853
Method:                 Least Squares   F-statistic:                     857.5
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          2.44e-183
Time:                        10:59:49   Log-Likelihood:                -940.21
No. Observations:                 444   AIC:                             1888.
Df Residuals:                     440   BIC:                             1905.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0476      0.097      0.489      0.625      -0.144       0.239
Mkt-RF         1.0705      0.022     48.037      0.000       1.027       1.114
SMB           -0.1398      0.033     -4.185      0.000      -0.206      -0.074
HML           -0.2176      0.033     -6.624      0.000      -0.282      -0.153
==============================================================================
Omnibus:                       18.531   Durbin-Watson:                   2.150
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               44.123
Skew:                           0.108   Prob(JB):                     2.62e-10
Kurtosis:                       4.529   Cond. No.                         4.74
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
Carhart
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.860
Model:                            OLS   Adj. R-squared:                  0.859
Method:                 Least Squares   F-statistic:                     676.3
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          3.94e-186
Time:                        10:59:49   Log-Likelihood:                -930.20
No. Observations:                 444   AIC:                             1870.
```

```
Df Residuals:                     439   BIC:                             1891.
Df Model:                           4
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.1242      0.097      1.283      0.200      -0.066       0.314
Mkt-RF         1.0440      0.023     46.194      0.000       1.000       1.088
SMB           -0.1408      0.033     -4.304      0.000      -0.205      -0.077
HML           -0.2561      0.033     -7.695      0.000      -0.322      -0.191
MOM           -0.0995      0.022     -4.500      0.000      -0.143      -0.056
==============================================================================
Omnibus:                       18.601   Durbin-Watson:                   2.114
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               39.714
Skew:                           0.182   Prob(JB):                     2.38e-09
Kurtosis:                       4.419   Cond. No.                         5.19
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
FF5
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.862
Model:                            OLS   Adj. R-squared:                  0.860
Method:                 Least Squares   F-statistic:                     545.1
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          1.74e-185
Time:                        10:59:49   Log-Likelihood:                -928.33
No. Observations:                 444   AIC:                             1869.
Df Residuals:                     438   BIC:                             1893.
Df Model:                           5
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.1651      0.099      1.667      0.096      -0.030       0.360
Mkt-RF         1.0239      0.024     42.925      0.000       0.977       1.071
SMB           -0.1558      0.036     -4.370      0.000      -0.226      -0.086
HML           -0.0747      0.043     -1.720      0.086      -0.160       0.011
RMW           -0.0839      0.047     -1.800      0.073      -0.176       0.008
CMA           -0.3162      0.066     -4.757      0.000      -0.447      -0.186
==============================================================================
Omnibus:                       17.214   Durbin-Watson:                   2.149
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               33.321
Skew:                           0.201   Prob(JB):                     5.81e-08
Kurtosis:                       4.280   Cond. No.                         5.30
==============================================================================
```

```
Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

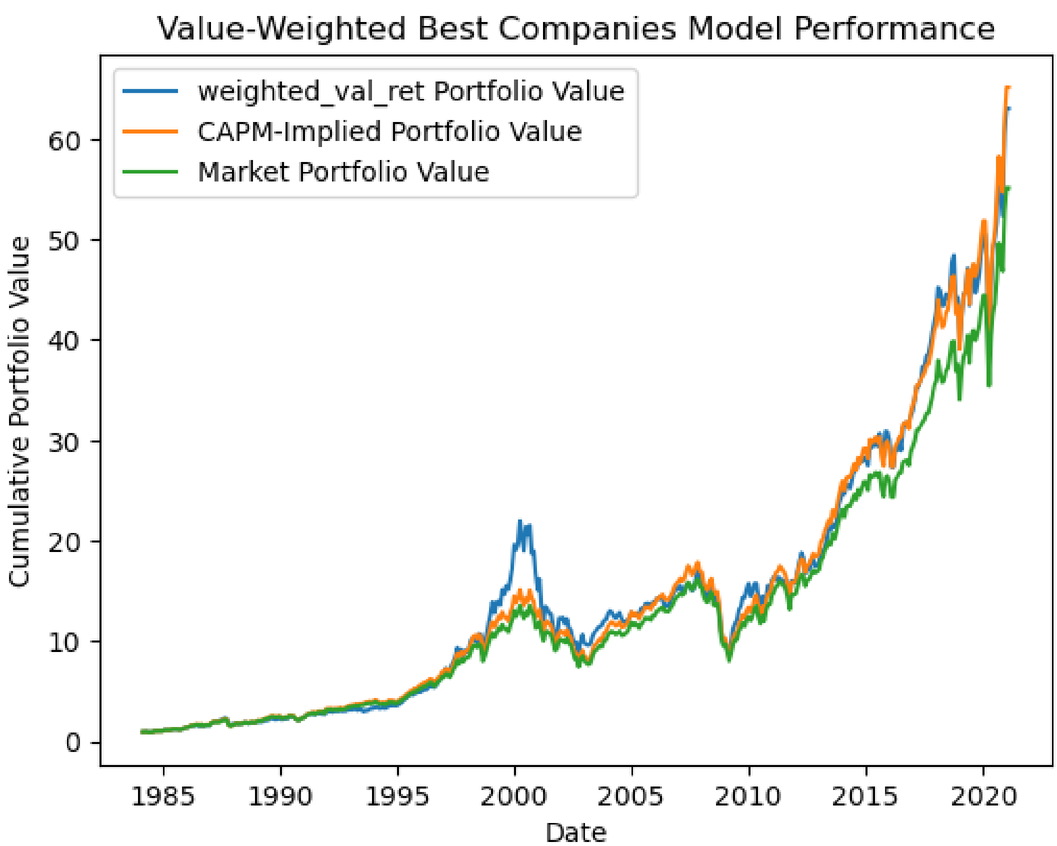
The equal-weighted produces positive alphas, but the alphas are insignificant except under the Carhart model, where they are significant at the 1% level. The value-weighted portfolios produce positive insignificant alphas, except under the FF5 model, where they are significant at the 10% level.

1.3.3 Part C

[12]:
```
val_ff5 = pd.merge(val_returns.to_frame(), ff5, how='inner', on=['date'])
assert len(val_ff5) == len(val_returns)

val_ff5['date'] = (val_ff5['date'].str[:4] + val_ff5['date'].str[5:]).
↪astype('int')

plot_cum_returns(val_ff5, 'weighted_val_ret', val_capm_beta, 'Value-Weighted␣
↪Best Companies')
```



The value-weighted Best Companies strategy is highly correlated to the market portfolio, which makes sense given its beta of 1.02. It performs almost the exact same as the CAPM model implied returns. Interestingly, the strategy outperforms the CAPM model in the late 1990's, but in the early 2000's loses all of those gains and remains roughly the same as the CAPM model.

1.3.4 Part D

[13]:
```
eq_returns_pre_jan1 = eq_returns[eq_returns.index < '2010-01']
eq_returns_post_jan1 = eq_returns[eq_returns.index >= '2010-01']

val_returns_pre_jan1 = val_returns[val_returns.index < '2010-01']
val_returns_post_jan1 = val_returns[val_returns.index >= '2010-01']
```

[14]:
```
estimate_models(eq_returns_pre_jan1, 'weighted_eq_ret', ff5_mom)
```

```
CAPM
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.894
Model:                            OLS   Adj. R-squared:                  0.893
Method:                 Least Squares   F-statistic:                     2599.
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          1.84e-152
Time:                        10:59:49   Log-Likelihood:                -624.62
No. Observations:                 311   AIC:                             1253.
Df Residuals:                     309   BIC:                             1261.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.1791      0.103      1.733      0.084      -0.024       0.382
Mkt-RF         1.1456      0.022     50.976      0.000       1.101       1.190
==============================================================================
Omnibus:                       59.444   Durbin-Watson:                   1.821
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              141.583
Skew:                           0.924   Prob(JB):                     1.80e-31
Kurtosis:                       5.741   Cond. No.                         4.63
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
FF3
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.904
Model:                            OLS   Adj. R-squared:                  0.903
Method:                 Least Squares   F-statistic:                     958.8
```

```
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          1.62e-155
Time:                        10:59:49   Log-Likelihood:                -609.52
No. Observations:                 311   AIC:                             1227.
Df Residuals:                     307   BIC:                             1242.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.1526      0.100      1.529      0.127      -0.044       0.349
Mkt-RF         1.1378      0.023     50.282      0.000       1.093       1.182
SMB            0.1791      0.033      5.508      0.000       0.115       0.243
HML            0.0612      0.034      1.804      0.072      -0.006       0.128
==============================================================================
Omnibus:                       57.788   Durbin-Watson:                   1.917
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              134.842
Skew:                           0.907   Prob(JB):                     5.24e-30
Kurtosis:                       5.668   Cond. No.                         4.88
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
Carhart
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.919
Model:                            OLS   Adj. R-squared:                  0.918
Method:                 Least Squares   F-statistic:                     872.8
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          6.41e-166
Time:                        10:59:49   Log-Likelihood:                -581.59
No. Observations:                 311   AIC:                             1173.
Df Residuals:                     306   BIC:                             1192.
Df Model:                           4
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.2899      0.093      3.115      0.002       0.107       0.473
Mkt-RF         1.0944      0.021     51.007      0.000       1.052       1.137
SMB            0.1800      0.030      6.045      0.000       0.121       0.239
HML            0.0041      0.032      0.127      0.899      -0.059       0.067
MOM           -0.1495      0.019     -7.760      0.000      -0.187      -0.112
==============================================================================
Omnibus:                       22.918   Durbin-Watson:                   1.952
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               42.241
Skew:                           0.427   Prob(JB):                     6.72e-10
Kurtosis:                       4.591   Cond. No.                         5.44
```

```
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
FF5
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.906
Model:                            OLS   Adj. R-squared:                  0.904
Method:                 Least Squares   F-statistic:                     588.2
Date:                Fri, 26 Apr 2024   Prob (F-statistic):         2.93e-154
Time:                        10:59:49   Log-Likelihood:                -605.47
No. Observations:                 311   AIC:                             1223.
Df Residuals:                     305   BIC:                             1245.
Df Model:                           5
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.1107      0.103      1.071      0.285      -0.093       0.314
Mkt-RF         1.1443      0.025     45.291      0.000       1.095       1.194
SMB            0.2177      0.035      6.184      0.000       0.148       0.287
HML            0.0601      0.046      1.308      0.192      -0.030       0.151
RMW            0.1124      0.045      2.485      0.013       0.023       0.201
CMA           -0.0601      0.067     -0.893      0.373      -0.193       0.072
==============================================================================
Omnibus:                       47.746   Durbin-Watson:                   1.915
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               95.887
Skew:                           0.809   Prob(JB):                     1.51e-21
Kurtosis:                       5.186   Cond. No.                         5.63
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

[14]: 1.1456255645790132

[15]:
```
estimate_models(val_returns_pre_jan1, 'weighted_val_ret', ff5_mom)
```

```
CAPM
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.823
Model:                            OLS   Adj. R-squared:                  0.822
Method:                 Least Squares   F-statistic:                     1436.
Date:                Fri, 26 Apr 2024   Prob (F-statistic):         3.57e-118
```

```
Time:                        10:59:49   Log-Likelihood:                -696.81
No. Observations:                 311   AIC:                             1398.
Df Residuals:                     309   BIC:                             1405.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0763      0.130      0.585      0.559      -0.180       0.333
Mkt-RF         1.0740      0.028     37.890      0.000       1.018       1.130
==============================================================================
Omnibus:                       22.393   Durbin-Watson:                   2.081
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               78.564
Skew:                          -0.051   Prob(JB):                     8.71e-18
Kurtosis:                       5.460   Cond. No.                         4.63
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
FF3
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.857
Model:                            OLS   Adj. R-squared:                  0.856
Method:                 Least Squares   F-statistic:                     613.9
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          2.50e-129
Time:                        10:59:49   Log-Likelihood:                -663.40
No. Observations:                 311   AIC:                             1335.
Df Residuals:                     307   BIC:                             1350.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.2108      0.119      1.777      0.077      -0.023       0.444
Mkt-RF         1.0325      0.027     38.371      0.000       0.980       1.085
SMB           -0.1836      0.039     -4.749      0.000      -0.260      -0.108
HML           -0.3140      0.040     -7.778      0.000      -0.393      -0.235
==============================================================================
Omnibus:                       18.769   Durbin-Watson:                   2.069
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               43.759
Skew:                           0.245   Prob(JB):                     3.15e-10
Kurtosis:                       4.771   Cond. No.                         4.88
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
```

```
specified.
Carhart
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.866
Model:                            OLS   Adj. R-squared:                  0.864
Method:                 Least Squares   F-statistic:                     493.7
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          4.50e-132
Time:                        10:59:49   Log-Likelihood:                -653.62
No. Observations:                 311   AIC:                             1317.
Df Residuals:                     306   BIC:                             1336.
Df Model:                           4
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.3103      0.117      2.645      0.009       0.079       0.541
Mkt-RF         1.0011      0.027     37.012      0.000       0.948       1.054
SMB           -0.1830      0.038     -4.876      0.000      -0.257      -0.109
HML           -0.3554      0.040     -8.825      0.000      -0.435      -0.276
MOM           -0.1082      0.024     -4.457      0.000      -0.156      -0.060
==============================================================================
Omnibus:                       17.969   Durbin-Watson:                   2.034
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               32.544
Skew:                           0.334   Prob(JB):                     8.57e-08
Kurtosis:                       4.437   Cond. No.                         5.44
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
FF5
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.862
Model:                            OLS   Adj. R-squared:                  0.859
Method:                 Least Squares   F-statistic:                     379.5
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          1.33e-128
Time:                        10:59:49   Log-Likelihood:                -658.55
No. Observations:                 311   AIC:                             1329.
Df Residuals:                     305   BIC:                             1352.
Df Model:                           5
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.3123      0.123      2.546      0.011       0.071       0.554
Mkt-RF         0.9899      0.030     33.031      0.000       0.931       1.049
```

```
SMB            -0.2018      0.042     -4.834      0.000      -0.284      -0.120
HML            -0.1985      0.055     -3.642      0.000      -0.306      -0.091
RMW            -0.0888      0.054     -1.655      0.099      -0.194       0.017
CMA            -0.2309      0.080     -2.891      0.004      -0.388      -0.074
==============================================================================
Omnibus:                       18.423   Durbin-Watson:                   2.054
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               36.181
Skew:                           0.312   Prob(JB):                     1.39e-08
Kurtosis:                       4.550   Cond. No.                         5.63
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

[15]: 1.07400676934913

[16]:
```python
estimate_models(eq_returns_post_jan1, 'weighted_eq_ret', ff5_mom)
```

```
CAPM
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.888
Model:                            OLS   Adj. R-squared:                  0.888
Method:                 Least Squares   F-statistic:                     1043.
Date:                Fri, 26 Apr 2024   Prob (F-statistic):           2.98e-64
Time:                        10:59:49   Log-Likelihood:                -259.57
No. Observations:                 133   AIC:                             523.1
Df Residuals:                     131   BIC:                             528.9
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0552      0.154     -0.358      0.721      -0.361       0.250
Mkt-RF         1.1356      0.035     32.303      0.000       1.066       1.205
==============================================================================
Omnibus:                       12.646   Durbin-Watson:                   2.328
Prob(Omnibus):                  0.002   Jarque-Bera (JB):               28.717
Skew:                          -0.293   Prob(JB):                     5.81e-07
Kurtosis:                       5.200   Cond. No.                         4.58
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
FF3
                            OLS Regression Results
```

```
==============================================================================
Dep. Variable:                      y   R-squared:                       0.898
Model:                            OLS   Adj. R-squared:                  0.895
Method:                 Least Squares   F-statistic:                     377.2
Date:                Fri, 26 Apr 2024   Prob (F-statistic):           1.22e-63
Time:                        10:59:49   Log-Likelihood:                -253.85
No. Observations:                 133   AIC:                             515.7
Df Residuals:                     129   BIC:                             527.3
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0314      0.153     -0.206      0.837      -0.333       0.271
Mkt-RF         1.0845      0.038     28.728      0.000       1.010       1.159
SMB            0.2192      0.065      3.397      0.001       0.092       0.347
HML           -0.0465      0.058     -0.798      0.426      -0.162       0.069
==============================================================================
Omnibus:                       10.516   Durbin-Watson:                   2.414
Prob(Omnibus):                  0.005   Jarque-Bera (JB):               25.784
Skew:                          -0.073   Prob(JB):                     2.52e-06
Kurtosis:                       5.152   Cond. No.                         4.95
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
Carhart
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.898
Model:                            OLS   Adj. R-squared:                  0.895
Method:                 Least Squares   F-statistic:                     281.5
Date:                Fri, 26 Apr 2024   Prob (F-statistic):           2.16e-62
Time:                        10:59:49   Log-Likelihood:                -253.67
No. Observations:                 133   AIC:                             517.3
Df Residuals:                     128   BIC:                             531.8
Df Model:                           4
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0230      0.154     -0.150      0.881      -0.327       0.281
Mkt-RF         1.0798      0.039     27.913      0.000       1.003       1.156
SMB            0.2169      0.065      3.346      0.001       0.089       0.345
HML           -0.0597      0.063     -0.954      0.342      -0.184       0.064
MOM           -0.0284      0.048     -0.589      0.557      -0.124       0.067
==============================================================================
```

```
Omnibus:                       10.864   Durbin-Watson:                   2.415
Prob(Omnibus):                  0.004   Jarque-Bera (JB):               27.365
Skew:                          -0.080   Prob(JB):                     1.14e-06
Kurtosis:                       5.216   Cond. No.                         5.26
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
FF5
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.916
Model:                            OLS   Adj. R-squared:                  0.913
Method:                 Least Squares   F-statistic:                     276.7
Date:                Fri, 26 Apr 2024   Prob (F-statistic):           1.78e-66
Time:                        10:59:49   Log-Likelihood:                -240.77
No. Observations:                 133   AIC:                             493.5
Df Residuals:                     127   BIC:                             510.9
Df Model:                           5
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0539      0.141      0.382      0.703      -0.225       0.333
Mkt-RF         1.0504      0.035     29.920      0.000       0.981       1.120
SMB            0.2615      0.064      4.111      0.000       0.136       0.387
HML            0.1160      0.063      1.827      0.070      -0.010       0.242
RMW            0.1714      0.092      1.868      0.064      -0.010       0.353
CMA           -0.5059      0.104     -4.876      0.000      -0.711      -0.301
==============================================================================
Omnibus:                        5.555   Durbin-Watson:                   2.399
Prob(Omnibus):                  0.062   Jarque-Bera (JB):                5.417
Skew:                           0.350   Prob(JB):                       0.0666
Kurtosis:                       3.698   Cond. No.                         5.14
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

[16]: 1.135639057256818

[17]:
```
estimate_models(val_returns_post_jan1, 'weighted_val_ret', ff5_mom)
```

```
CAPM
                            OLS Regression Results
==============================================================================
```

```
Dep. Variable:                      y   R-squared:                       0.866
Model:                            OLS   Adj. R-squared:                  0.865
Method:                 Least Squares   F-statistic:                     844.6
Date:                Fri, 26 Apr 2024   Prob (F-statistic):           5.71e-59
Time:                        10:59:49   Log-Likelihood:                -266.55
No. Observations:                 133   AIC:                             537.1
Df Residuals:                     131   BIC:                             542.9
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.1346      0.163     -0.827      0.410      -0.457       0.187
Mkt-RF         1.0767      0.037     29.062      0.000       1.003       1.150
==============================================================================
Omnibus:                        0.491   Durbin-Watson:                   2.275
Prob(Omnibus):                  0.782   Jarque-Bera (JB):                0.462
Skew:                           0.140   Prob(JB):                        0.794
Kurtosis:                       2.933   Cond. No.                         4.58
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
FF3
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.870
Model:                            OLS   Adj. R-squared:                  0.867
Method:                 Least Squares   F-statistic:                     288.3
Date:                Fri, 26 Apr 2024   Prob (F-statistic):           5.42e-57
Time:                        10:59:49   Log-Likelihood:                -264.29
No. Observations:                 133   AIC:                             536.6
Df Residuals:                     129   BIC:                             548.1
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.1888      0.165     -1.144      0.255      -0.515       0.138
Mkt-RF         1.1140      0.041     27.281      0.000       1.033       1.195
SMB           -0.1162      0.070     -1.665      0.098      -0.254       0.022
HML           -0.0469      0.063     -0.744      0.458      -0.172       0.078
==============================================================================
Omnibus:                        0.648   Durbin-Watson:                   2.284
Prob(Omnibus):                  0.723   Jarque-Bera (JB):                0.487
Skew:                           0.148   Prob(JB):                        0.784
Kurtosis:                       3.024   Cond. No.                         4.95
```

```
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
Carhart
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.872
Model:                            OLS   Adj. R-squared:                  0.868
Method:                 Least Squares   F-statistic:                     217.6
Date:                Fri, 26 Apr 2024   Prob (F-statistic):           4.59e-56
Time:                        10:59:49   Log-Likelihood:                -263.47
No. Observations:                 133   AIC:                             536.9
Df Residuals:                     128   BIC:                             551.4
Df Model:                           4
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.1694      0.165     -1.024      0.308      -0.497       0.158
Mkt-RF         1.1032      0.042     26.490      0.000       1.021       1.186
SMB           -0.1216      0.070     -1.742      0.084      -0.260       0.016
HML           -0.0771      0.067     -1.145      0.254      -0.210       0.056
MOM           -0.0650      0.052     -1.255      0.212      -0.168       0.038
==============================================================================
Omnibus:                        0.459   Durbin-Watson:                   2.267
Prob(Omnibus):                  0.795   Jarque-Bera (JB):                0.327
Skew:                           0.121   Prob(JB):                        0.849
Kurtosis:                       3.014   Cond. No.                         5.26
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
FF5
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.892
Model:                            OLS   Adj. R-squared:                  0.888
Method:                 Least Squares   F-statistic:                     210.3
Date:                Fri, 26 Apr 2024   Prob (F-statistic):           1.22e-59
Time:                        10:59:49   Log-Likelihood:                -251.93
No. Observations:                 133   AIC:                             515.9
Df Residuals:                     127   BIC:                             533.2
Df Model:                           5
Covariance Type:            nonrobust
==============================================================================
```

```
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0739      0.153     -0.482      0.631      -0.377       0.230
Mkt-RF         1.0790      0.038     28.261      0.000       1.003       1.155
SMB           -0.1331      0.069     -1.924      0.057      -0.270       0.004
HML            0.1444      0.069      2.091      0.038       0.008       0.281
RMW           -0.0546      0.100     -0.547      0.585      -0.252       0.143
CMA           -0.5724      0.113     -5.072      0.000      -0.796      -0.349
==============================================================================
Omnibus:                        0.982   Durbin-Watson:                   2.447
Prob(Omnibus):                  0.612   Jarque-Bera (JB):                0.614
Skew:                           0.136   Prob(JB):                        0.736
Kurtosis:                       3.191   Cond. No.                         5.14
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

[17]: 1.0767224921961434

The strategy worked well before January 1st, 2010. The alphas were positive for the equal and value weighted strategies, and statistically significant at the 1% level when tested using the Carhart model. The alphas and their significance levels, as well as the factor coefficients, are extremely similar to Table 3 in the Edmands paper.

The strategy does not work in the post-period, and has negative alphas under both strategies and all models with the exception of the equal weighted portfolio under FF5, which is positive and insignificant. In commentary on Table 4, the Edmands paper argues that higher alphas under the Feb 1998 - Dec 2009 period, after Forbes began publishing the Best Companies list every year, are evidence that reasons other than lack of public information are behind this market misevaluation. The authors suggest the difficulty of incorporating this information into a traditional evaluation model as one possible reason. Perhaps what happened post Jan 2010 is that with increasing competition in the hedge fund industry and advances in automatic sentiment analysis and online information (ex, Problem 3d), evaluation models got more advanced and were able price employee sentiment more effectively.

1.3.5 Part E

[18]:
```
'''
Load Industry portfolio data from Ken French's website. Using monthly␣
↪value-weighted industry portfolios.
'''

industries = pd.read_csv('12_Industry_Portfolios.CSV')
industries = industries.astype({'date': 'string'})
industries['date'] = industries['date'].apply(lambda x: x[:4] + '-' + x[4:])
industries
```

```python
val_returns_pre_1999 = val_returns[val_returns.index < '1999-01']
val_returns_post_1999 = val_returns[val_returns.index >= '1999-01']

df_pre = pd.merge(industries, val_returns_pre_1999, on=["date"], how="inner")
df_pre = pd.merge(df_pre, ff5_mom, on=["date"], how="inner")
assert len(df_pre) == len(val_returns_pre_1999)

df_post = pd.merge(industries, val_returns_post_1999, on=["date"], how="inner")
df_post = pd.merge(df_post, ff5_mom, on=["date"], how="inner")
assert len(df_post) == len(val_returns_post_1999)

print('Pre-1999')
print(sm.OLS(df_pre['weighted_val_ret'] - df_pre['RF'], sm.
 ↪add_constant(df_pre[['NoDur','Durbl','Manuf','Enrgy','Chems','BusEq','Telcm','Utils','Shops
 ↪fit().summary())

print('Post-1999')
print(sm.OLS(df_post['weighted_val_ret'] - df_post['RF'], sm.
 ↪add_constant(df_post[['NoDur','Durbl','Manuf','Enrgy','Chems','BusEq','Telcm','Utils','Shop
 ↪fit().summary())
```

```
Pre-1999
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.932
Model:                            OLS   Adj. R-squared:                  0.927
Method:                 Least Squares   F-statistic:                     189.5
Date:                Fri, 26 Apr 2024   Prob (F-statistic):           3.62e-90
Time:                        10:59:49   Log-Likelihood:                -291.15
No. Observations:                 179   AIC:                             608.3
Df Residuals:                     166   BIC:                             649.7
Df Model:                          12
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.4601      0.105     -4.366      0.000      -0.668      -0.252
NoDur          0.1690      0.053      3.162      0.002       0.063       0.275
Durbl          0.0158      0.035      0.447      0.655      -0.054       0.085
Manuf          0.1394      0.082      1.708      0.090      -0.022       0.301
Enrgy          0.1215      0.028      4.354      0.000       0.066       0.177
Chems          0.1616      0.056      2.898      0.004       0.052       0.272
BusEq          0.4446      0.029     15.467      0.000       0.388       0.501
Telcm          0.0604      0.034      1.764      0.080      -0.007       0.128
Utils         -0.0594      0.041     -1.457      0.147      -0.140       0.021
Shops          0.0132      0.045      0.290      0.772      -0.077       0.103
```

```
Hlth             0.1453      0.037      3.933      0.000       0.072       0.218
Money            0.0855      0.041      2.094      0.038       0.005       0.166
Other           -0.3550      0.066     -5.388      0.000      -0.485      -0.225
==============================================================================
Omnibus:                        9.461   Durbin-Watson:                   2.200
Prob(Omnibus):                  0.009   Jarque-Bera (JB):               10.952
Skew:                           0.406   Prob(JB):                      0.00419
Kurtosis:                       3.899   Cond. No.                         17.7
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
Post-1999
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.900
Model:                            OLS   Adj. R-squared:                  0.896
Method:                 Least Squares   F-statistic:                     189.6
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          1.22e-118
Time:                        10:59:49   Log-Likelihood:                -526.70
No. Observations:                 265   AIC:                             1079.
Df Residuals:                     252   BIC:                             1126.
Df Model:                          12
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const           -0.1958      0.117     -1.674      0.095      -0.426       0.035
NoDur            0.1078      0.058      1.857      0.064      -0.007       0.222
Durbl            0.0565      0.024      2.378      0.018       0.010       0.103
Manuf            0.0386      0.059      0.652      0.515      -0.078       0.155
Enrgy            0.0409      0.024      1.731      0.085      -0.006       0.087
Chems            0.0097      0.055      0.176      0.861      -0.099       0.119
BusEq            0.5496      0.029     19.191      0.000       0.493       0.606
Telcm            0.0069      0.036      0.192      0.848      -0.064       0.077
Utils           -0.1259      0.035     -3.571      0.000      -0.195      -0.056
Shops           -0.0223      0.047     -0.477      0.634      -0.115       0.070
Hlth             0.1356      0.038      3.600      0.000       0.061       0.210
Money            0.1066      0.039      2.750      0.006       0.030       0.183
Other            0.0089      0.070      0.127      0.899      -0.130       0.147
==============================================================================
Omnibus:                        8.068   Durbin-Watson:                   2.239
Prob(Omnibus):                  0.018   Jarque-Bera (JB):               12.893
Skew:                          -0.133   Prob(JB):                      0.00159
Kurtosis:                       4.047   Cond. No.                         17.0
==============================================================================
```

```
Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

Yes, based on the regression it appears that the composition of the best companies to work for has somewhat changed over time. Prior to Jan 1, 1999, the value weighted returns had statistically significant (1%) and sizable positive coefficients (>0.1) with the NoDur, Enrgy, Chems, BusEq, and Hlth industries. After Jan 1, 1999, the value weighted returns had statistically significant (1%) and sizable positive coefficients (>0.1) with the BusEQ, Utils, Hlth, and Money industries. This indicates that after Jan 1 1999, the Utils and Money industries became better to work for on average, as they became more correlated with our portfolio, while the NoDur, Enrgy and Chems industries got a little worse and the BusEq and Hlth industries remained good to work in, on average.

1.4 Problem 3

1.4.1 Part A

At around 1.1 for equal-weighted and 1.05 for value-weighted under all four models, the beta of this strategy is very close to 1. Our strategy is long a bundle of specifically selected market assets, which we expect to have a beta of approximately 1 across the board, since we haven't shorted anything or made any explicit selections based on beta value.

With the right weights, a long-short portfolio could be beta neutral, making it more hedged against market risk. This would make it attractive to institutional investors, who are looking for high alpha funds. At the same time, shorting the market portfolio with positive beta would significantly decrease the expected returns of the strategy, making it less attractive to retail investors. It would be interesting to compare the Sharpe Ratio of this new long-short portfolio.

1.4.2 Part B

As we can see from the results, estimating the Carhart model from the long portfolio on the best companies to work with (the companies with the highest employee satisfaction) produces a positive alpha that is significant at the 1% level, for both the equal-weighted and value-weighted portfolios. The other three (CAPM, FF3, FF5) models also produce positive alphas, with some significant and some insignificant. Since we were able to generate alpha by simply buying and holding companies that have high employee satisfaction, this shows that financial markets do not fully price the value of employee satisfaction.

1.4.3 Part C

One possible reason for why the alpha decreases over time is that people became more aware of this employee satisfaction strategy, and with improvements in automatic sentiment analysis, this became easier to implement at a larger scale (ex, Problem 3d). As more people become aware of the performance of this strategy and start incorporating it more effectively into their own strategies, the efficiencies of the market grow, allowing for financial markets to price the value of employee satisfaction more. This would cause the alpha to decrease over time, decreasing the performance of this strategy.

1.4.4 Part D

Among the many ways in which the employee satisfaction strategy can be improved for the modern world by tracking what employees are reading about in real time, two stand out to us:

Firstly, we can observe the proportion of the time that employees are reading about anything unrelated to their work in general to estimate how distracted employees of a particular company are. Distracted employees would likely be a good indicator of employee disinterest and dissatisfaction and vice versa. Even if distracted employees do not mean that they are disinterested or dissatisfied, it cannot be good for a company's productivity and performance if many of their employees are distracted on the job.

Secondly, we can observe the frequency with which employees browse specific content that indicate disatisfaction with their job, or about life in general. Employees browsing content related to finding a new job, dealing with their authoritarian boss, improving their depression, etc. with higher than average frequency is a good indication that they are dassatisfied and that their company will underperform.

[ ]:
```

```