Python Backtesting System

02 Feb 2022

Sector Rotation Strategy – HK

Interview Candidate:

Andy Chan andyinter1@gmail.com +852 9603 3105

Content

Introduction of the back-test python system	. 3
Strategy Backtesting	11

Please refer to the python system in below link for more details https://github.com/ccfandy1/Backtesting---Sector-Rotation

Introduction of the back-test python system

The system aims to test the feasibility of sector rotation investment strategy. The strategy will use growth of basic earnings per share as an indicator of outperforming sectors. Throughout of the testing period, the backtesting will generate semiannually investment signals and weights based on sector growth and market cap respectively. Certain numbers of outperforming sectors will be invested, and the virtual portfolio performance calculated afterwards will eventually be compared with the benchmark index performance for strategy validation.

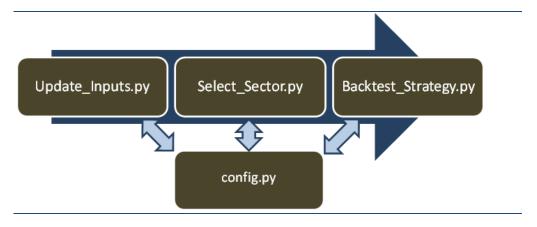
Strategy components:

Indicator	Growth of basic earnings per share
Signal	Sectors with high sector growth.
	For simplification, investment signal directly triggers investment actions without other considerations
Rule	Trade entry and exit follow the same signal-based rule
Initial parameter	 Trade days in a year: 252 Objective function: annualized sharpe ratio Virtual portfolio initial NAV: 1,000,000 *please refer to configuration table in Strategy Backtesting section for test specific parameters
Initial data	 Stocks related: earnings growth, stock price, market cap Benchmark related: market index value

Python program structure summary

The backtesting system contains 4 modules: 1) Update_Inputs, 2) Select_Sector, 3) Backtest_Strategy, and 4) Config:

Figure 1 Python structure



Brief introduction of 4 modules:

Update_Inputs	To cleanse Bloomberg data and update input CSV accordingly
Select_Sector	 Calculate sector growth and rank them Generate investment signals based on ranking
Backtest_Strategy	Calculate weighted price growth of each stock in each chosen sector, then calculate portfolio return performance accordingly Perform backtesting, show backtest stats and relevant graphs afterwards
Config	Contains all global variable, most of independent variable, and utility methods for other modules to use during execution

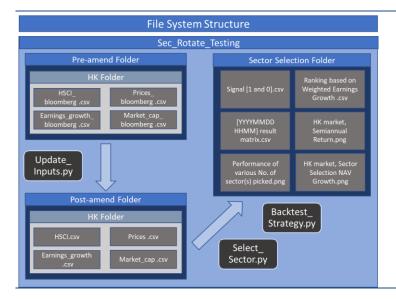
The system is written with object-oriented programming structure and every CSV file has its own class in the program.

Update_Inputs module can be executed individually to update inputs, while Select_Sector and Backtest_Strategy modules need inputs from Update_Inputs module. Backtest_Strategy needs additional inputs from Select_Sector module as well.

Most of independent variables are in the config module where user can interact with the program. For example, user can alter a specific set of testing parameters, such as time period, in config module to perform various forms of backtesting.

File System Structure

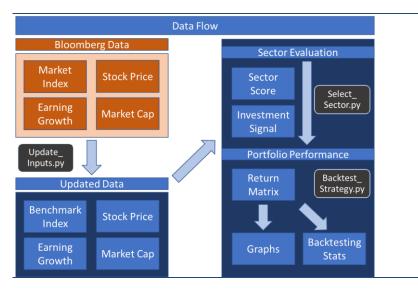
Figure 2 File system structure overview



There are 3 folders containing all system input and output, namely 1) "pre-amend", 2) "post-amend", and 3) "sector selection", which are placed under "sec_rotate_testing" parent folder. Raw and cleansed input CSV are stored in "pre-amend" and "post-amend" folder respectively, while the signal data and final testing outcomes are stored in "sector selection" folder. (fig. 2)

Execution flow

Figure 3 Execution flow overview



Bloomberg Data

Four sets of data will be extracted from Bloomberg to csv, including: 1) market index, 2) earnings growth, 3) stock price, and 4) market cap. The first set (fig. 4) is for calculating benchmark performance while the latter three (fig. 5 & 6), which contain data of chosen stocks, are for calculating performance of sector rotation strategy.

Figure 4 Example of bloomberg raw data (HSCI_bloomberg.csv)



Figure 5 Example of bloomberg raw data (earnings_growth_bloomberg.csv)

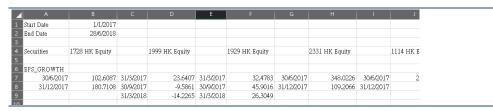
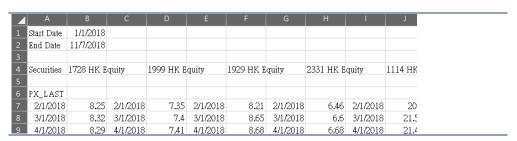


Figure 6 Example of bloomberg raw data (prices.csv)



Updated Data - Update_Inputs.py

Inputs	 HSCI_bloomberg.csv (fig. 4) – raw market index data Earnings_growth_bloomberg.csv (fig. 5) – raw earnings growth data of stocks market cap_bloomberg.csv – raw market cap data of stocks prices_bloomberg.csv (fig. 6) – raw price data of stocks
Outcomes	 HSCI.csv – contains date index, and daily closed price data of specific index from 2000 to 2018. Earnings growth.csv (fig. 7) – contains date index, stock code and sector header, and quarterly weighted earnings growth data (earnings growth / total earnings growth of stock's sector) of chosen stocks from around 1985 to 2018. Prices.csv (fig. 8) – contains date index, stock code header, and daily closed trading price data of chosen stocks from 2000 to 2018. Market cap.csv – contains date index, stock code and sector header, and quarterly market cap data of chosen stocks from around 1985 to 2018.

This module will check if any new data point is found in the raw Bloomberg data. New data will be cleansed, standardized, and updated to its corresponding input csv file accordingly (fig. 7 & 8) for further processing of other modules.

Mainly, the historical data of market index, stock prices, market cap, and earnings growth will be used to calculate benchmark KPI, price return of stock candidates, investment weights of stock candidates, and sector growth respectively.

Figure 7 Example of cleansed data (earnings growth.csv)

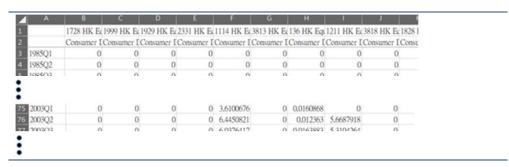
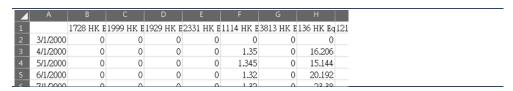


Figure 8 Example of cleansed data (prices.csv)



Sector Evaluation – Select_Sector.py

Inputs	 Earnings growth.csv (fig. 7) – cleansed stock earnings growth data Market cap.csv – cleansed stock market cap data
Outcomes	 Ranking based on weighted earnings growth.csv (fig. 9) – contains date index, sector header, and ranking data of user specified period. Signal [1 and 0].csv (fig. 10) – contains date index, sector header, and investment signal data of user specified period.

This module will generate investment signals for backtesting simulation.

Since not every HK listed stock will disclose its quarterly data, time unit of data record is semi-annual base. Semiannually, earnings growth of stocks under each sector will be accumulated and become sector growth. Sectors will then be ranked and scored based on

their growth, while a certain number of outperforming sectors, which can be specified by user manually, will be picked and generate investment signals accordingly.

Cleansed earnings growth data is the main input to be used in this module. Sector score and investment signal results will be exported into csv called "Ranking based on Weighted Earnings Growth.csv" and "Signal [1 and 0].csv" respectively (fig. 9 & 10).

It should be noted that "1" and "0" shown in investment signal csv (fig. 10) respectively means "good to invest" and "skip the sector".

Figure 9 Sector Ranking (Ranking based on weighted earnings growth.csv)

Г		Consumer	Consumer	Energy	Financials	Health Car	Industrials	Information	Materials	Real Estate	Telecomm	Utilities
2 1	.998Q4	8	3	0	2	10	6	1	7	9	5	4
3 1	.999Q2	5	0	6	9	10	7	3	4	2	1	8
4 1	999Q4	7	0	8	9	10	6	2	5	1	3	4
5 2	2000Q2	8	7	9	5	10	4	0	2	3	1	6
6 2	2000Q4	5	2	6	4	8	10	1	3	9	0	7

Figure 10 Investment Signals (Signal [1 and 0].csv)

Consumer Consumer Consumer Energy Financials Health Car/ Industrials Information Materials Real Estate Telecomm\(Utilities 1999\) Consumer Consum			Consumer	Concurac	Energy	Financiale	Health Car	Industrials	Informatio	Materiale	Paal Fetate	Telecommi	Hilitiae
4 1999Q4 0 1 0 0 0 0 1 0 1 1	2	1998Q4	0	1	bliefgy 1	. 1	0	0	1	0	Cear Estate	0	Cunnes
	3	1999Q2	0	1	(0	0	0	1	1	1	1	(
5 2000Q2 0 0 0 0 0 1 1 1 1 1	4	1999Q4	0	1	(0	0	0	1	0	1	1	
	5	2000Q2	0	0	(0	0	1	1	1	1	1	(
6 2000Q4 0 1 0 1 0 0 1 1 0 1	6	2000Q4	0	1	(1	0	0	1	1	0	1	(

Portfolio Performance - Backtest_Strategy.py

Inputs	 Signal [1 and 0].csv – contains semiannually investment signal presented as "1" and "0" Market cap.csv –cleansed quarterly stock market cap data Stock prices.csv – cleansed daily stock price data Market Index.csv – cleansed benchmark market index data
Outcomes	 [YYYYMMDD HHMM] result matrix.csv – contains weighted price return of each invested stock and returns of virtual portfolio Backtesting stats – show backtest KPI, such as drawdown Performance graphs – show benchmark comparison and semi-annual return of investment strategy

This module will be responsible for generating virtual portfolio performance and conducting backtest.

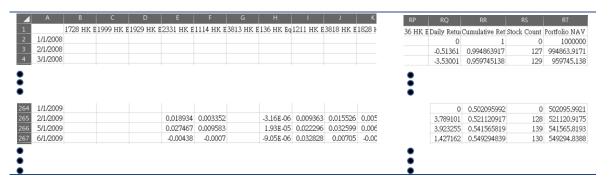
Part 1 – Virtual portfolio performance

Semiannually, based on sector investment signals, market cap of stocks under chosen sectors will be picked to calculate their investment weights, which will be multiplied by each daily stock price return in next period to get weighted investment return. For example, daily price return on any trading date in 1H22 will multiply its investment weight in 2H21, and so on.

By accumulating all weighted investment returns on each trading date, daily return of virtual portfolio will be calculated, which will be used to calculate other KPI such as cumulative return and portfolio NAV over time.

Investment signals, stock price data and market cap data are the main inputs for portfolio return calculation. Portfolio return data generated will be exported into csv called "[YYYYMMDD HHMM] result matrix.csv". (fig. 11)

Figure 11 Return Matrix



Part 2 - Backtesting

The second part of this module is to conduct backtesting on the sector selection strategy. There are two types of testing methods can be chosen from: 1) single holdout cross validation, and 2) rolling window testing. User can change testing parameters in config module to alter the tests.

Main test parameters a user can set are successively testing method, beginning year of testing period, in-sample period, out-of-sample period, minimum and maximum sector count to be tested, time interval between each rolling window and rolling frequency (fig. 12). Time unit of all numbers are year.

Figure 12 Test parameters in config module for single holdout cross validation

```
test_method = 'rolling'
test_beg_yr = 2009
train_set_period = 7
test_set_period = 3
optim_param = 'sharpe'
sim_sec_cnt_min = 3
sim_sec_cnt_max = 8
roll_increm_yr = 0
roll_freq = 1
```

For example, the setting shown in figure 12 can be read as "10 years testing period since 2009, of which 7 years in-sample and 3 years out-of-sample. Perform backtest 1 time, with 0 year time interval between each run. Try sector count from 3 to 8 with in-sample data and test the performance of the one that can maximize sharpe ratio with out-of-sample data."

Under "rolling" test method, setting roll_increm_year to 0 and roll_freq to 1 for single rolling window validation holdout cross validation test (fig. 12), while setting roll_increm_year to >0 and roll_freq >1 for rolling window testing (fig. 13).

Figure 13 Test parameters in config module for rolling window

```
test_method = 'rolling'
test_beg_yr = 2009
train_set_period = 2
test_set_period = 1
optim_param = 'sharpe'
sim_sec_cnt_min = 3
sim_sec_cnt_max = 8
roll_increm_yr = 1
roll_freq = 5
```

On the other hand, "single period" method does not separate data into in-sample and outof-sample. Instead of finding optimal trading parameters from the training data set, it directly uses user manual input of trading parameters in the whole testing period. Under this method, roll_increm_year and roll_freq will be set to 0 and 1 respectively (fig. 14).

Figure 14 Test parameters in config module for single period test method

```
test_method = 'single period'
test_beg_yr = 2016
train_set_period = 0
test_set_period = 3
optim_param = 'sharpe'
sim_sec_cnt_min = 3
sim_sec_cnt_max = 8
roll_increm_yr = 0
roll_freq = 1
```

For example, after running rolling window test, user may find that several versions of trading parameters are acceptable for the strategy. User can then conduct his/her further evaluation and judgement, manually input final decision, and test it with another set of data. Further details will be elaborated in the next section.

Portfolio return matrix and market index data are the main inputs for backtesting. Backtesting results such as annual sharpe ratio will be shown in python console while related graphs will be exported in PNG format. (fig. 15 - 16)

Figure 15 Outputs of single holdout cross validation: in-sample test

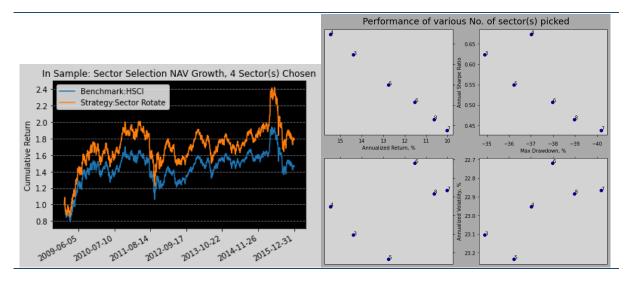


Figure 16 Outputs of single holdout cross validation: out-of-sample test

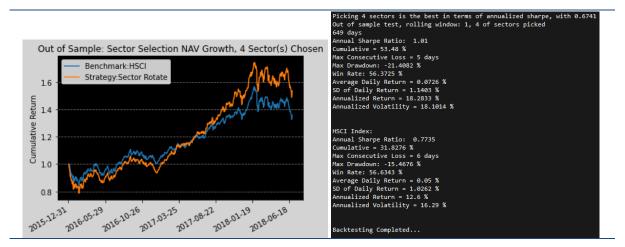
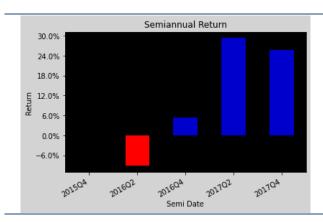


Figure 17 Outputs of single holdout cross validation: out-of-sample test



For single holdout cross validation and test method of "single period", an additional semiannual return graph will be shown and exported as well. (fig. 17)

Strategy Backtesting

Single holdout cross validation

Over the 10 years after 2008, investment environment is characterized by record-low nominal interest rate and widespread usage of QE. A single holdout cross validation test will be conducted on these 10 years data.

For simplification purpose, we assume statistical stationarity during this period of 'new normal' and thus no regime analysis is needed to conduct accordingly. Transaction costs and dividend income are not considered during testing.

Backtest configuration:

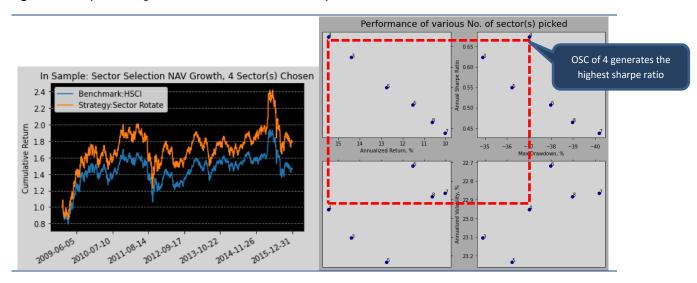
Testing market	Hong Kong
Benchmark	HSCI index
Evaluation metrics	Annual sharpe ratio (to be maximized)
Time period	10 years, from Jan 2009 to Jan 2019:
	 Due to data accessibility, only contain data up to 2Q18 Using data after year 2008 to avoid distortion from 2008 financial crisis. Using data before year 2019 to avoid distortion from HK unrest and COVID-19
IS/OOS split ratio	2/3 data to be in-sample, remaining 1/3 to be out-of-sample
Testing sector count	3 to 8 sectors, out of total 11 sectors
	to diversify idiosyncratic risk, must pick at least 3 sectors

Figure 18 Test configuration

```
test_method = 'rolling'
test_beg_yr = 2009
train_set_period = 7
test_set_period = 3
optim_param = 'sharpe'
sim_sec_cnt_min = 3
sim_sec_cnt_max = 8
roll_increm_yr = 0
roll_freq = 1
```

2:1 ratio is applied to determine in-sample and out-of-sample data. As a result, 10 years of data after year 2018 will be separated into two segments: first 7 years are in-sample while remaining $^{\sim}3$ years are out-of-sample. Based on system simulation on the in-sample data, the optimal sector count (hereafter 'OSC'), which generate the highest annual sharpe ratio, is 4. (fig. 19)

Figure 19 Outputs of single holdout cross validation: in-sample test



The system then refreshes investment signals based on OSC parameter and tests the strategy again on the second segment of data. Based on the backtest stats (fig. 20), the rotation strategy performance outperforms HSCI benchmark performance in terms of annual sharpe ratio (1.01 vs 0.774) and cumulative return (53.5% vs 31.83%).

Figure 20 Outputs of single holdout cross validation: out-of-sample test

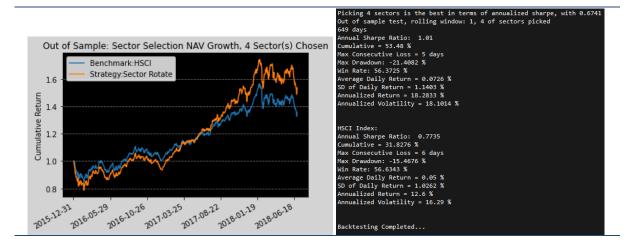
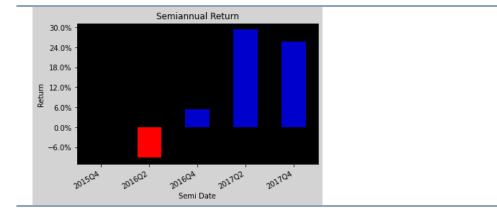


Figure 21 Outputs of single holdout cross validation: out-of-sample test



Walk Forward Optimization - Optimal Sector Count

To reduce potential overfitting, rolling window test will be conducted on the first segment of testing data to further optimize OSC.

Given 7 years of data availability (from Jan 2009 to Dec 2015), we set rolling window period as 3 years and testing frequency as 5 times. 2:1 ratio is again applied, splitting each 3 years testing window into 2 years in-sample and 1 year out-of-sample. (fig. 22)

Additional configuration:

Rolling increment	1 year (next window slides 1 year)
Rolling frequency	5 times (5 windows)

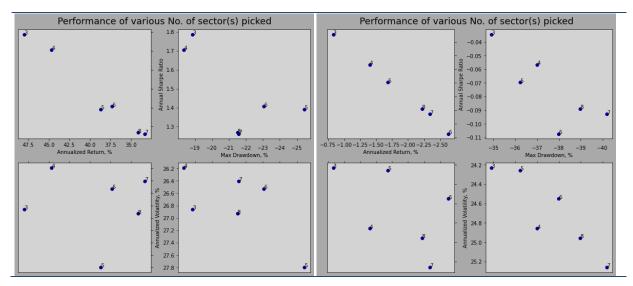
Figure 22 Test configuration

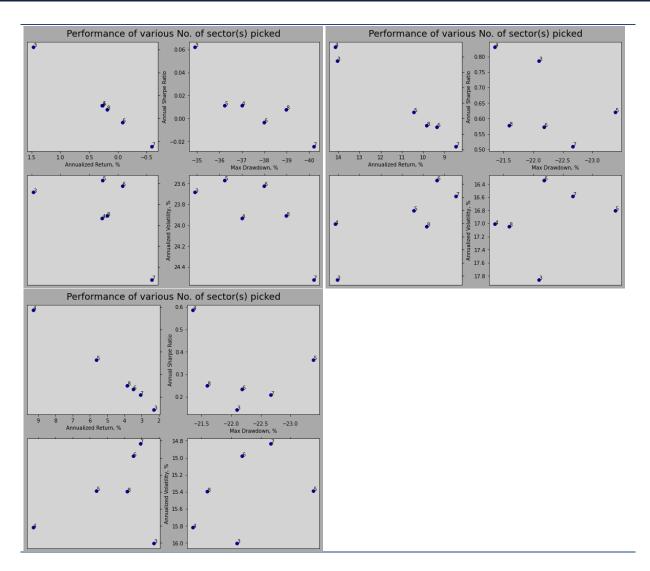
```
test_beg_yr = 2009
train_set_period = 2
test_set_period = 1
optim_param = 'sharpe'
sim_sec_cnt_min = 3
sim_sec_cnt_max = 8
roll_increm_yr = 1
roll_freq = 5
```

In-sample tests - optimal sector count filtering

In 5 in-sample test runs, OSC of 3 and 4 perform well in terms of annual sharpe ratio (fig. 23). As a result, both OSC are further evaluated in their respective 1 year out-of-sample test.

Figure 23 In-sample test KPI comparison (time window 1 to 5, from left to right)





Out-of-sample tests - optimal sector count testing

Based on out-of-sample results of 5 test runs (3 tests with OSC of 3; 2 tests with OSC of 4), sector rotation strategy with OSC of 4 beats benchmark in terms of annual sharpe ratio in both tests while OSC of 3 loses to benchmark in one of the three tests (fig. 24). Given that other KPI such as win rate do not offer strong evidence to differentiate one OSC's performance from the other, OSC of 4 will be picked considering that it offers higher diversification against idiosyncratic risk of sector.

Figure 24 Out-of-sample test KPI comparison

```
Picking 3 sectors is the best in terms of annualized sharpe, with 1.7875
Out of sample test, rolling window: 1, 3 of sectors picked
                                                                                                                                                                                                         Picking 3 sectors is the best in terms of annualized sharpe,
Out of sample test, rolling window: 2, 3 of sectors picked
260 days
Annual Sharpe Ratio:
                                                                                                                                                                                                           261 days
                                                                                                                                                                                                           Annual Sharpe Ratio: :
Cumulative = 22.6391 %
                                                                                                                                                                                                                                                                    1.141
Cumulative = -20.6664 %
Max Consecutive Loss = 7 days
Max Drawdown: -34.9069 %
Win Rate: 47.7551 %
                                                                                                                                                                                                         Max Consecutive Loss = 8 days
Max Drawdown: -20.3443 %
Win Rate: 51.8219 %
Average Daily Return = 0.0852 %
SD of Daily Return = 1.1855 %
Annualized Return = 21.4722 %
Win Mate: 47.7551 %
Average Daily Return = -0.0738 %
SD of Daily Return = 1.7451 %
Annualized Return = -18.6007 %
Annualized Volatility = 27.702 %
                                                                                                                                                                                                          Annualized Volatility = 18.8185 %
                                                                                                                                                                                                          HSCI Index:
                                                                                                                                                                                                          Annual Sharpe Ratio: 1.1849
Cumulative = 19.6805 %
Annual Sharpe Ratio: -0.8591
Cumulative = -22.9207 %
Cumulative = -22.9207 %
Max Consecutive Loss = 8 days
Max Drawdown: -36.0094 %
Win Rate: 48.1633 %
Average Daily Return = -0.0918 %
SD of Daily Return = 1.6968 %
Annualized Return = -23.1417 %
Annualized Volatility = 26.9363 %
                                                                                                                                                                                                          Max Consecutive Loss = 8 days
Max Drawdown: -17.1014 %
Win Rate: 53.8776 %
                                                                                                                                                                                                         Average Daily Return = 0.0786 %
SD of Daily Return = 1.0527 %
Annualized Return = 19.8007 %
Annualized Volatility = 16.7111 %
```

```
Picking 3 sectors is the best in terms of annualized sharpe, with 0.0622
Out of sample test, rolling window: 3, 3 of sectors picked
                                                                                                                                                                          Picking 4 sectors is the best in terms of annualized sharpe, with 0.8316
Out of sample test, rolling window: 4, 4 of sectors picked
 261 days
Annual Sharpe Ratio: 0.3934
                                                                                                                                                                           261 days
Annual Sharpe Ratio: 0.7755
 Manual sharpe Railu: 0.3934
Cumulative = 5.5543 %
Max Consecutive Loss = 10 days
Max Drawdown: -22.0978 %
Win Rate: 49.5902 %
Average Daily Return = 0.0263 %
                                                                                                                                                                           Manual Sharpe Ratio: 0.7735
Cumulative = 11.6665 %
Max Consecutive Loss = 5 days
Max Drawdown: -10.2435 %
Win Rate: 53.4413 %
                                                                                                                                                                           Average Daily Return = 0.0469 % SD of Daily Return = 0.9596 % Annualized Return = 11.8131 %
 SD of Daily Return = 1.0629 %
Annualized Return = 6.6374 %
 Annualized Volatility = 16.8734 %
                                                                                                                                                                            Annualized Volatility = 15.2336 %
 HSCI Index:
                                                                                                                                                                           HSCI Index:
ASCL Index:
Annual Sharpe Ratio: 0.1937
Cumulative = 1.7475 %
Max Consecutive Loss = 8 days
Max Drawdown: -17.2208 %
Win Rate: 48.1481 %
                                                                                                                                                                             Annual Sharpe Ratio: 0.0742
                                                                                                                                                                           Cumulative = 0.0861 %
Max Consecutive Loss = 7 days
                                                                                                                                                                            Max Drawdown: -9.3662 %
Win Rate: 52.8455 %
Average Daily Return = 0.0117 %
SD of Daily Return = 0.9628 %
Annualized Return = 2.9608 %
Annualized Volatility = 15.2845 %
                                                                                                                                                                           MIN Rate: 52.045p &
Average Daily Return = 0.004 %
SD of Daily Return = 0.8545 %
Annualized Return = 1.0061 %
Annualized Volatility = 13.564 %
 Picking 4 sectors is the best in terms of annualized sharpe, with 0.5869
Out of sample test, rolling window: 5, 4 of sectors picked
Out of Sample
261 days
Annual Sharpe Ratio: -0.0773
Cumulative = -5.3346 %
Max Consecutive Loss = 8 days
Max Drawdown: -33.7907 %
Win Rate: 48.583 %
Aleeage Daily Return = -0.0079
 Average Daily Return = -0.0079 % SD of Daily Return = 1.622 % Annualized Return = -1.9916 %
 Annualized Volatility = 25.7483 %
 HSCI Index:
 Annual Sharpe Ratio: -0.3118
Cumulative = -8.6612 %
Max Consecutive Loss = 9 days
Max Drawdown: -29.8616 %
Win Rate: 47.9675 %
  Average Daily Return = -0.0272 %
 SD of Daily Return = 1.3866 %
Annualized Return = -6.8637 %
 Annualized Volatility = 22.011 %
```

Test on second segment of data

Once the optimal OSC is determined, sector rotation strategy is now tested with the second segment of data from 2016 to 2019, and test configuration is accordingly changed (fig. 25). It should be noted that the OSC picked is manually inputted in "num of sec chosen" variable.

Figure 25 Test configuration

```
test_method = 'single period'
test_beg_yr = 2016
train_set_period = 0
test_set_period = 3
optim_param = 'sharpe'
sim_sec_cnt_min = 3
sim_sec_cnt_max = 8
roll_increm_yr = 0
roll_freq = 1
num_of_sec_chosen = 4
```

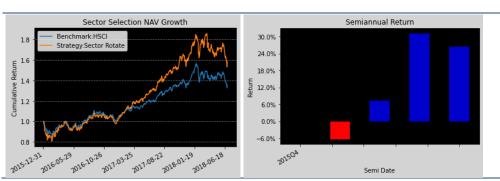
With the same testing period and OSC, the same result with single holdout cross validation is generated (fig. 26). The performance sector rotation strategy successfully outperforms the performance of HSCI index benchmark, with annual sharpe ratio of 1.053 vs benchmark's 0.855, and cumulative return of 58.22% vs benchmark's 35.62%, representing an alpha of 22.6%.

Figure 26 Test results

```
649 days
Annual Sharpe Ratio:
                      1.0534
Cumulative = 58.2207 %
Max Consecutive Loss = 6 days
Max Drawdown: -19.4795 %
Win Rate: 56.0458 %
Average Daily Return = 0.0776 %
SD of Daily Return = 1.1691 %
Annualized Return = 19.5493 %
Annualized Volatility = 18.5584 %
HSCI Index:
Annual Sharpe Ratio: 0.8553
Cumulative = 35.6247 %
Max Consecutive Loss = 6 days
Max Drawdown: -15.4676 %
Win Rate: 56.9558 %
Average Daily Return = 0.0551 %
SD of Daily Return = 1.0233 %
Annualized Return = 13.8935 %
Annualized Volatility = 16.2445 %
```

It should be noted that, the cumulative return graph (fig. 27) potentially indicate that the sector rotation strategy may be good at capturing the rising momentum as well, which can be further examined into in the future.

Figure 27 Semi return and cumulative return of the strategy



In conclusion, investors can benefit more from investing using sector rotation strategy than investing with passive index tracking strategy.