Python Backtesting System

14 Apr 2022

Sector Rotation Strategy – Hong Kong market

Interview Candidate:

Andy Chan andyinter1@gmail.com +852 9603 3105

Content

ntroduction of the back-test python system	3
Explanation of System Execution Flow	6
Strategy Backtesting	14

System Environment

External Libraries	Numpy, Pandas, PyQt5, SciPy, Matplotlib, itertools
Python Version	3.7

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 backtesting system aims to test the feasibility of sector rotation investment strategy. The strategy will use growth of basic earnings per share as an indicator to filter out outperforming sectors in specific market. Throughout of the testing period, the system 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. The system is written with object-oriented programming structure.

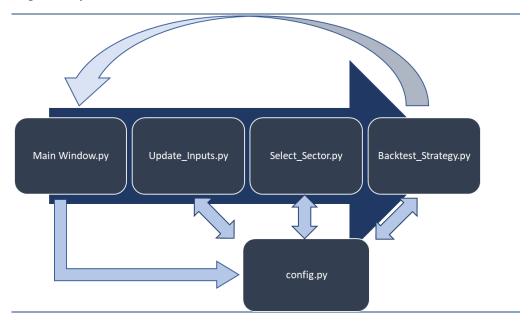
Strategy components:

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	Invest in outperforming sectors. Rebalance semi-annually
	Trade entry and exit follow the same signal-based rule
Initial data	 Stocks related: earnings growth stock price market cap Benchmark related:
	- market index value

Python program structure summary

The backtesting system contains 5 modules: 1) Main Window, 2) Update_Inputs, 3) Select_Sector, 4) Backtest_Strategy, and 5) Config (fig. 1). The Main Window module will transfer user inputs to config module. Then the remaining modules will conduct strategy backtest along with config module and feed the final outcomes back to Main Window module for display.

Figure 1 Python structure



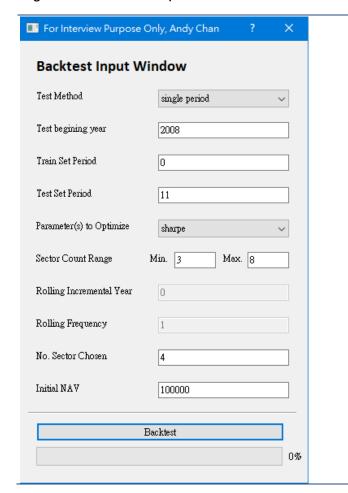
Brief introduction of 5 modules:

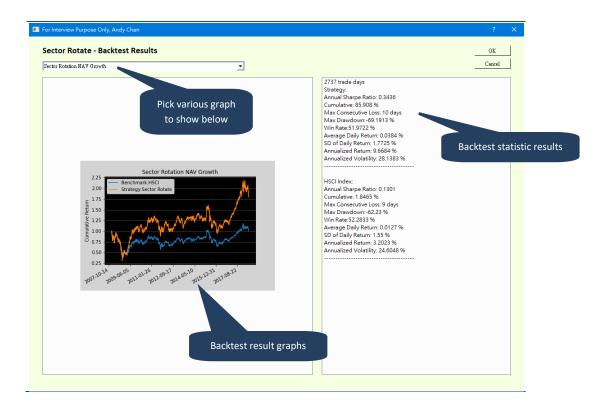
Main Window	To show input and result window forms
Update_Inputs	To cleanse Bloomberg data and update input CSV accordingly
Select_Sector	 Calculate sector growth and rank them Generate investment signals based on ranking
Backtest_Strategy	Calculate weighted price growth of each stock in each chosen sector, 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

Update_Inputs module can be executed individually to update inputs, while the updated inputs will be forwarded to Select_Sector module for further processing. Data from both Update_Inputs and Select_Sector module will then be forwarded to Backtest_Strategy module for final processing.

Most of input variables are in the config module, where user can interact with through window forms generated with Main Window module to perform various forms of backtesting. (fig. 2)

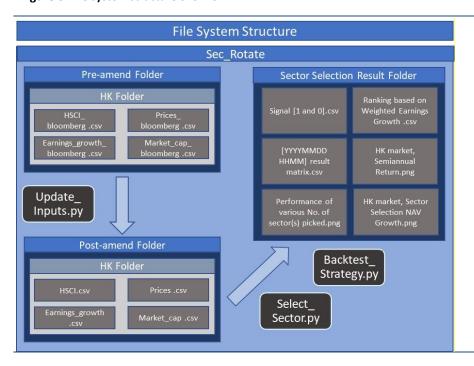
Figure 2 User Interface - Input Window & Result Window





File System Structure

Figure 3 File system structure overview

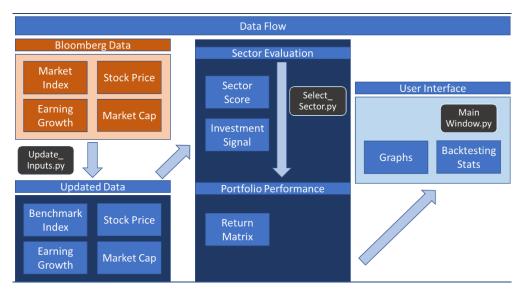


There are 3 folders containing all system input and output, namely 1) "pre-amend", 2) "post-amend", and 3) "sector selection", which are placed under "sec_rotate" parent folder. Raw and cleansed input CSV are stored in "pre-amend" and "post-amend" folder respectively, while the signal data and final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

Explanation of System Execution Flow

Execution flow

Figure 4 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. 5) is for calculating benchmark performance while the latter three (fig. 6, 7, 8), which contain data of chosen stocks, are for calculating performance of sector rotation strategy.

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

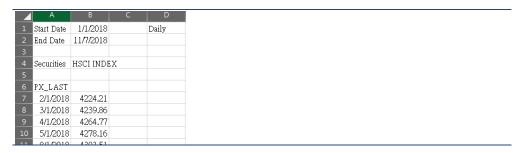


Figure 6 Example of bloomberg raw data (earnings_growth_bloomberg.csv)

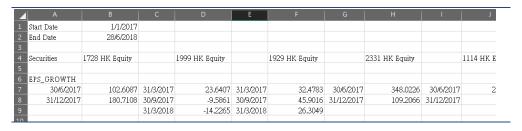


Figure 7 Example of bloomberg raw data (prices_bloomberg.csv)

1	А	В	С	D	Е	F	G	Н	l I	J
1	Start Date	1/1/2018								
2	End Date	11/7/2018								
3										
4	Securities	1728 HK E	quity	1999 HK E	quity	1929 HK E	quity	2331 HK E	quity	1114 HK
5										
6	PX_LAST									
7	2/1/2018	8.25	2/1/2018	7.35	2/1/2018	8.21	2/1/2018	6.46	2/1/2018	20
8	3/1/2018	8.32	3/1/2018	7.4	3/1/2018	8.65	3/1/2018	6.6	3/1/2018	21.5
9	4/1/2018	8.29	4/1/2018	7.41	4/1/2018	8.68	4/1/2018	6.68	4/1/2018	21.4

Figure 8 Example of bloomberg raw data (market cap_bloomberg.csv)

1	А	В	С	D	E	F	G	Н	I	J
1	Start Date	1/1/2018		Quarterly						
2	End Date	28/6/2018								
3										
4	Securities	1728 HK E	quity	1999 HK E	quity	1929 HK E	quity	2331 HK E	quity	1114 HK Ec
5										
6	HISTORIC	AL_MARK	ET_CAP							
7	30/3/2018	14201.77	30/3/2018	23891.58	30/3/2018	89400	30/3/2018	17522.92	30/3/2018	82641.51
8										

Updated Data - Update_Inputs.py

Inputs	HSCI_bloomberg.csv (fig. 5) – raw market index data
	Earnings_growth_bloomberg.csv (fig. 6) – raw earnings growth data of stocks
	market cap_bloomberg.csv (fig. 7) – raw market cap data of stocks
	• prices_bloomberg.csv (fig. 8) – 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. 9) – 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. 10) – 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 there is any new data point in the raw Bloomberg data. New data will be cleansed, standardized, and updated to its corresponding input csv file accordingly (fig. 9 & 10) 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 9 Example of cleansed data (earnings growth.csv)

A A	В		D		E		G	н		J	
	1728 HK I	1999 HK	Ec 1929 H	E(2331	HK E	1114 HK E	3813 HK E	136 HK Equ	1211 HK E	3818 HK E	1828
	Consumer	I Consume	Consum	er I Cons	umer I	Consumer I	Consumer 1	Consumer I	Consumer I	Consumer	I Const
1985Q1		0	0	0	0	0	0	0	0	()
4 1985Q2		0	0	0	0	0	0	0	0	()
1005/12		ni	n	n	0	-		0	n		
1005C/2 75 2003Q1		0	0	0	0	3.6100676	0	0.0160868	0	0	1
		0	0	0	0		0		0	0	

Figure 10 Example of cleansed data (prices.csv)

	Α	В	С	D	E	F	G		Н	
1		1728 HK E	1999 HK	E 1929 HK I	2331 HK E	1114 HK E	3813 HK	E 13	6 HK Eq	121
2	3/1/2000	0		0 0	0	0	()	0	
3	4/1/2000	0		0 0	0	1.35	()	16.206	
4	5/1/2000	0		0 0	0	1.345	()	15.144	
5	6/1/2000	0		0	0	1.32	()	20.192	
6	7/1/2000					1 20		١.	22 20	

Sector Evaluation - Select_Sector.py

Sector Lvaid	iation – select_sector.py
Inputs	Earnings growth.csv (fig. 9) – cleansed stock earnings growth data
	Market cap.csv – cleansed stock market cap data
Outcomes	 Ranking based on weighted earnings growth.csv (fig. 11) – contains date index, sector header, and ranking data of user specified period.
	 Signal [1 and 0].csv (fig. 12) – 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. 11 & 12).

It should be noted that "1" and "0" shown in investment signal csv (fig. 12) respectively means "good to invest" and "skip the sector".

Figure 11 Sector Ranking (Ranking based on weighted earnings growth.csv)

		Consumer	Consumer	Energy	Financials	Health Car	Industrials	Information	Materials	Real Estate	Telecomm	Utilities
1998	Q4	8	3	() 2	10	6	1	7	9	5	
1999	Q2	5	0	6	9	10	7	3	4	2	1	
1999	Q4	7	0	8	9	10	6	2	5	1	3	
2000	Q2	8	7	9	5	10	4	0	2	3	1	
2000	Q4	5	2	6	5 4	8	10	1	3	9	0	

Figure 12 Investment Signals (Signal [1 and 0].csv)

ď		Consumer	Consumer	Energy	Financiale	Health Can	Industrials	Informatio	Materiale	Real Estate	Telecommi	Litilities
2	1998Q4	0	1	I I	1	0	0	1	0	Cear Dotate	0	1
3	1999Q2	0	1	0	0	0	0	1	1	1	1	0
4	1999Q4	0	1	0	0	0	0	1	0	1	1	1
5	2000Q2	0	0	0	0	0	1	1	1	1	1	0
6	2000Q4	0	1	0	1	0	0	1	1	0	1	0

Portfolio Performance & User Interface - Backtest_Strategy.py & Main Window.py

Portfolio Performance & User Interface – Backtest_Strategy.py & Main Window.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.

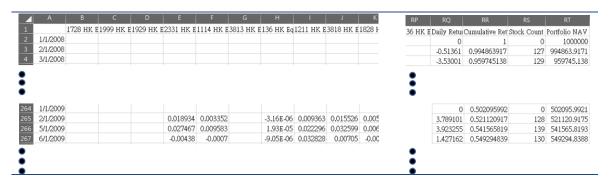
Part 1 – Virtual portfolio performance

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. 13)

Figure 13 Return Matrix

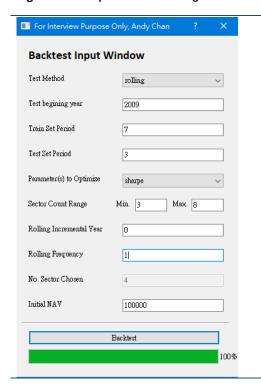


Part 2 - Backtesting

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 input window (fig. 14) to alter the test.

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, rolling frequency, number of sectors chosen, and initial NAV.

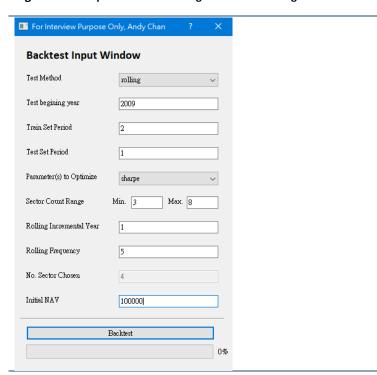
Figure 14 Test parameters for single holdout cross validation



For example, the setting shown in fig. 14 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."

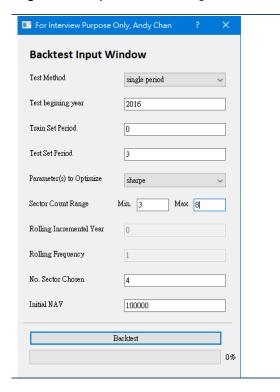
Under "rolling" test method, setting Rolling Incremental Year to 0 and Rolling Frequency to 1 for single holdout cross validation test (execute 1 time, no incremental year needed), while setting Rolling Incremental Year to > 0 and Rolling Frequency > 1 for rolling window testing (fig. 15).

Figure 15 Test parameters in config module for rolling window



On the other hand, "single period" method does not separate data into in-sample and out-of-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, Rolling Incremental Year and Rolling Frequency will be set to 0 and 1 respectively (fig. 16).

Figure 16 Test parameters in config module for single period test method



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 own evaluation and judgement, manually input final decision, and test it with another set of

data under single period test method. 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 result window while related graphs will also be exported to corresponding folder. (fig. 17 & 18)

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

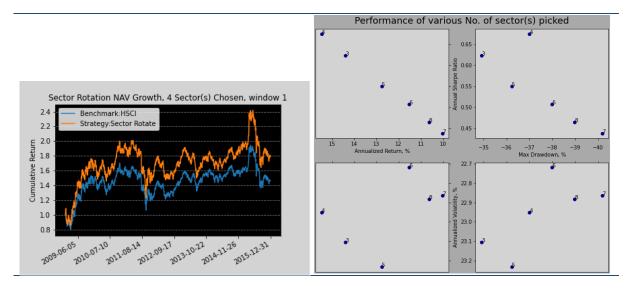


Figure 18 Outputs of single holdout cross validation: out-of-sample test

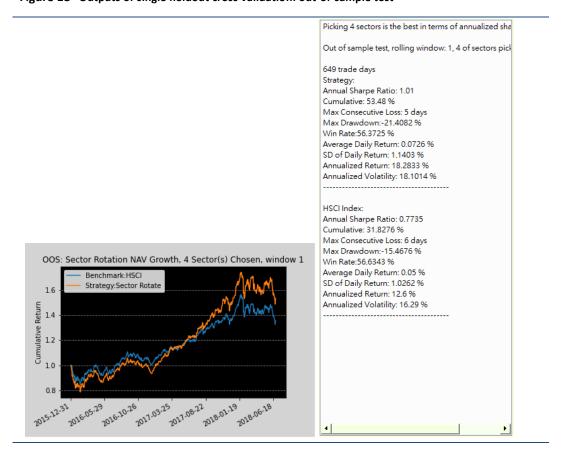
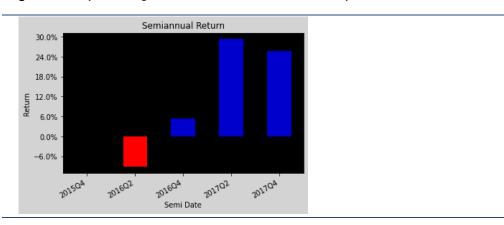


Figure 19 Outputs of single holdout cross validation: out-of-sample test



For single holdout cross validation and test method of "single period", an additional semiannual return graph will be shown and exported as well. (fig. 19)

Strategy Backtesting

Single holdout cross validation

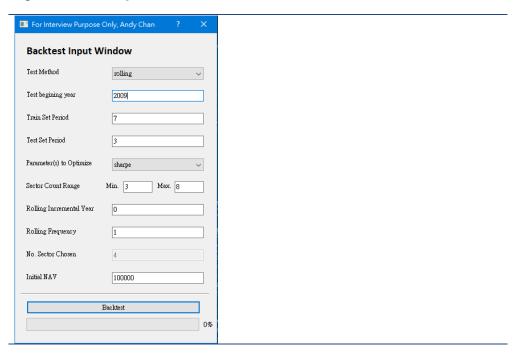
Over the 10 years after 2008, investment environment is characterized by record-low nominal interest rate and widespread usage of QE. A single holdout cross validation test will be conducted on these 10 years data.

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:

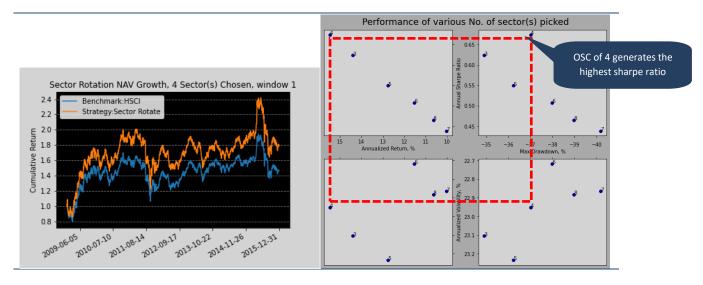
Testing market	Hong Kong
Benchmark	HSCI index
Evaluation metrics	Annual sharpe ratio (to be maximized)
Time period	10 years, from Jan 2009 to Jan 2019:
	 Due to data accessibility, only contain data up to 2Q18 Using data after year 2008 to avoid distortion from 2008 financial crisis. Using data before 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 20 Test configuration



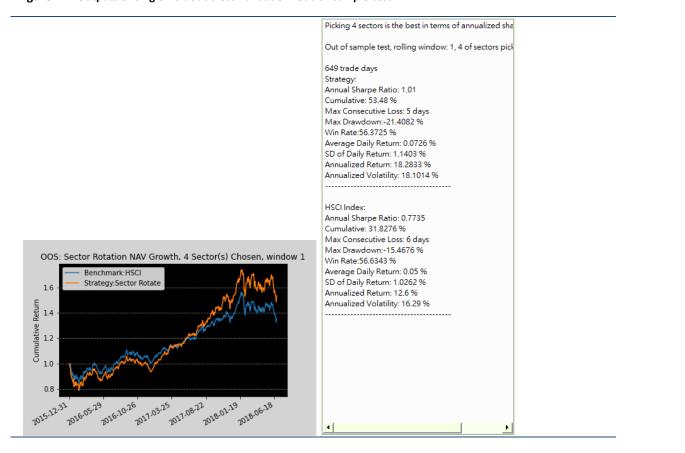
2:1 ratio is applied to determine in-sample and out-of-sample data. As a result, 10 years of data after year 2018 will be separated into two segments: first 7 years are in-sample while remaining ~3 years are out-of-sample (fig. 20). Based on system simulation on the in-sample data, the optimal sector count (hereafter 'OSC'), which generate the highest annual sharpe ratio, is 4. (fig. 21)

Figure 21 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. 22), 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 22 Outputs of single holdout cross validation: out-of-sample test



Semiannual Return

30.0% -24.0% -18.0% -12.0% -6.0% -0.0% --6.0% -2015QA 2016QA 2017QA 2017QA
Semi Date

Figure 23 Outputs of single holdout cross validation: out-of-sample test

Walk Forward Optimization - Optimal Sector Count

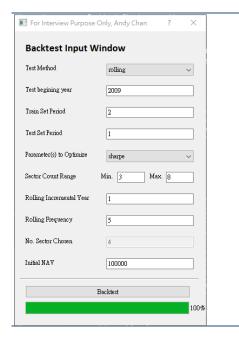
To reduce potential overfitting, rolling window test will be conducted on the first segment of testing data to further optimize OSC.

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. 24)

Additional configuration:

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

Figure 24 Test configuration



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. 25). As a result, both OSC are further evaluated in their respective 1 year out-of-sample test.

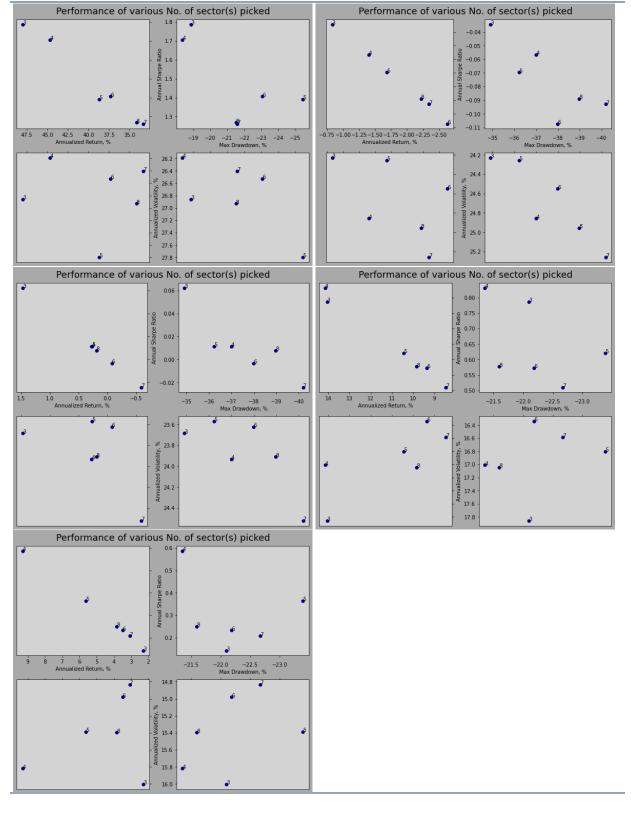
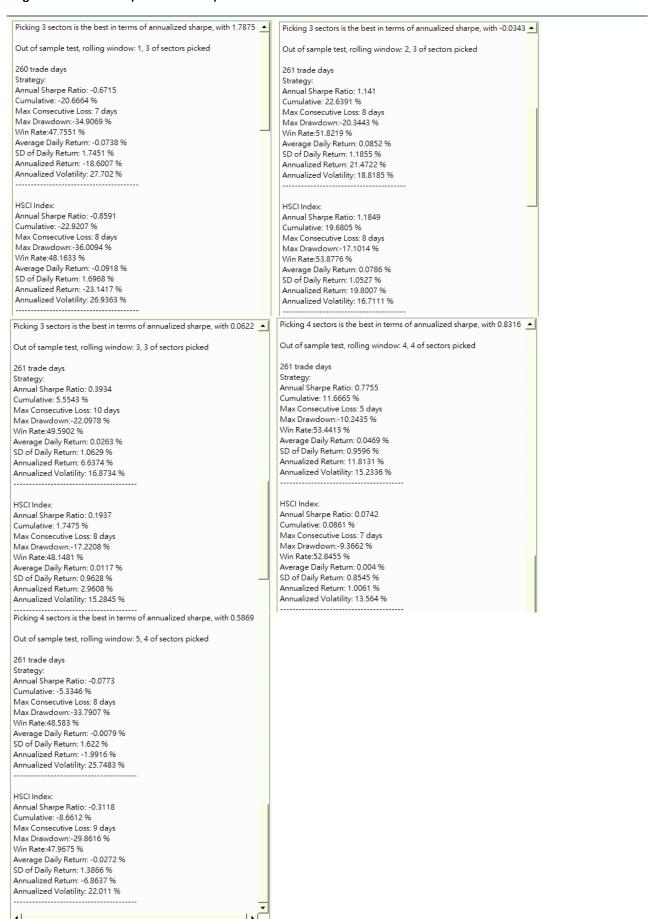


Figure 25 In-sample test KPI comparison (time window 1 to 5, from left to right)

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. 26). 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.

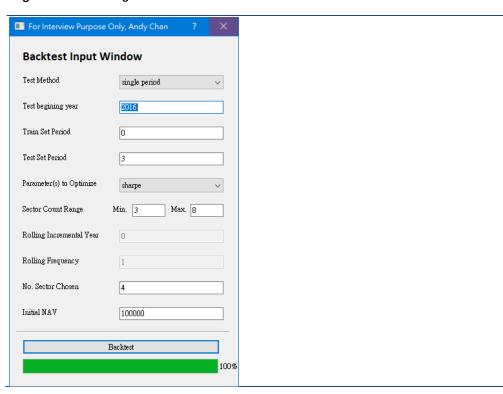
Figure 26 Out-of-sample test KPI comparison



Test on second segment of data

Once the optimal OSC is determined, user can then input it into No. Sector Chosen field and start testing with the second segment of data from 2016 to 2019.

Figure 27 Test configuration



With the same testing period and OSC, the same result with single holdout cross validation is generated (fig. 28). 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 28 Test results

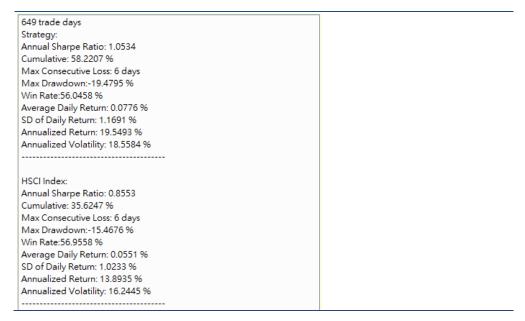
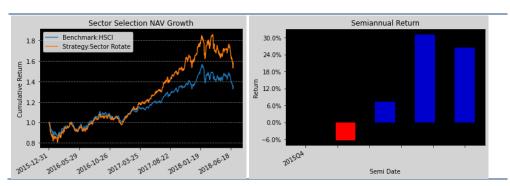


Figure 29 Semi return and cumulative return of the strategy



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