HFT Strategy based on MER and BO

By Shengsong Qu, Zicheng He

Mainly, there are two regimes in the market. One is the mean reversion state, which means that the price oscillates around the average level. The volume and participation remain balanced. The market movement is directionless and the volatility and the daily range of the price is small. In this state, we could buy at the lower point and sell at the higher point and repeat this process to make profit. The other is break-out state. It occurs when there exists imbalance in the volume traded, so the price would move out of the past period ranges. If we simply sell at the high point with the wrong assumption that the market is in the mean-reversion state we would lose money. Hence, it is necessary for us to use the imbalance between the traded volume to identify the break-out period.

The logic of our trading strategy is to first identify the market regimes. If we find that the market is in the mean-reversion state, we would use Bollinger bands method to generate signal for our trade. On the other hand, if we find that there exists great imbalance between the long and short position and the market is in the break-out state, we would use first flat our mean-reversion position and open our break-out position based on the volume. Specifically, if the position is smaller than a threshold, we would long the asset; while if the position is greater than another threshold, we would short the asset.

1. Alpha for Mean Reversion state.

1.2 Identification for the Mean Reversion state

In the mean reversion state, the price oscillates around the average level. The volume and participation remain balanced. The range of the price is small and remains at the same level, which means that there is an upper bond and a lower bond for the price. Based on this fact, we use rolling window test method to identify the mean reversion state. We set the lookback(test) period to 30 trading hours, because we find this is the appropriate time for the market to build up a mean reversion stage. The upper bond for the price is the N moving average of the price plus 2.0 * the N moving standard deviation of the price, and the lower bond for the price is the N moving average of the price minus the 2 * N moving standard deviation of the price. If we find that all the price in the test period window is greater than the lower bond and smaller than the upper bond and the volume is balanced, we would conclude that the market is in the mean reversion state, and we could use our mean-reversion strategy to trade. Otherwise, we could move forward our test window until we find the mean reversion period.

1.3 Strategy for the Mean Reversion state

After we find the mean reversion state, we would use 30 strategy period to create our mean-reversion strategy. We would use the Bollinger Band Method. A Bollinger Band is a technical analysis tool defined by a set of lines plotted two standard deviations (positively and negatively) away from a simple moving average (SMA) of the security's price, but can be adjusted to user preferences. Our moving average base line is the N trading hours moving average of the price.

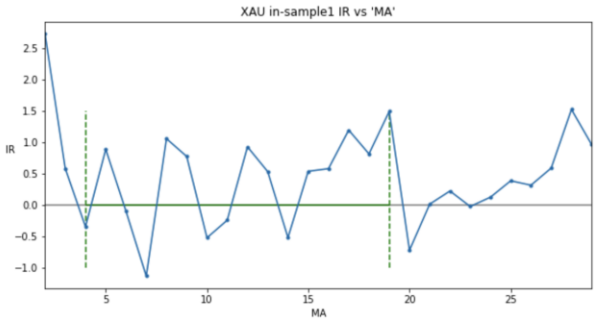
A Mean Reversion alpha is a signal “opposite” to a previous move or return – this market “move” can be taken the price difference between two previous dates.

MER signal = - diff(price some time ago vs price some other time ago)

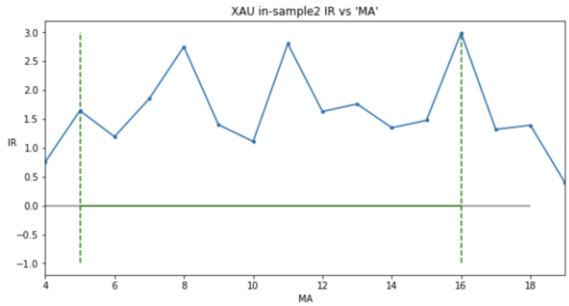
If the price fall below the moving average base line minus one rolling standard deviation, we would generate a long signal. When the price moves up to the moving average base line, we would flat the position to win the profit. If the price rise above the moving average base line plus one rolling standard deviation, we would generate a sell signal. When the price moves down to the moving average base line, we would flat the position to get the profit.

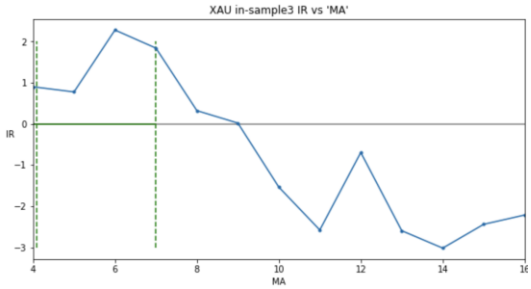
1.4 Test for a good parameter for moving average period in in-sample period

Just make it clear, first of all, we focus on the asset "XAU_USD", and work on all the process on this asset. And the training set is the first half of the data and the test set is the second half data.
The parameter N, which is the moving average period for the mean reversion Bollinger bands strategy, it is used both in the identification period and also the strategy period. Since we may need a N work relatively well on all period, so we may not optimize it on the whole in-sample period. Here is how we find it:
First, we divide the training data(in-sample data) into 3 subsets. Using the first one to run a grid search for N. The performance is judged by Information Ratio, which is the stat considered first in our research. Below shows the in-sample grid search of moving average N:



As we find it good for N between 4 and 19, we re-do a grid search within this range in the next 1/3 data of the training set, the data is called in- sample2. Same on rest of the data, called in-sample3.

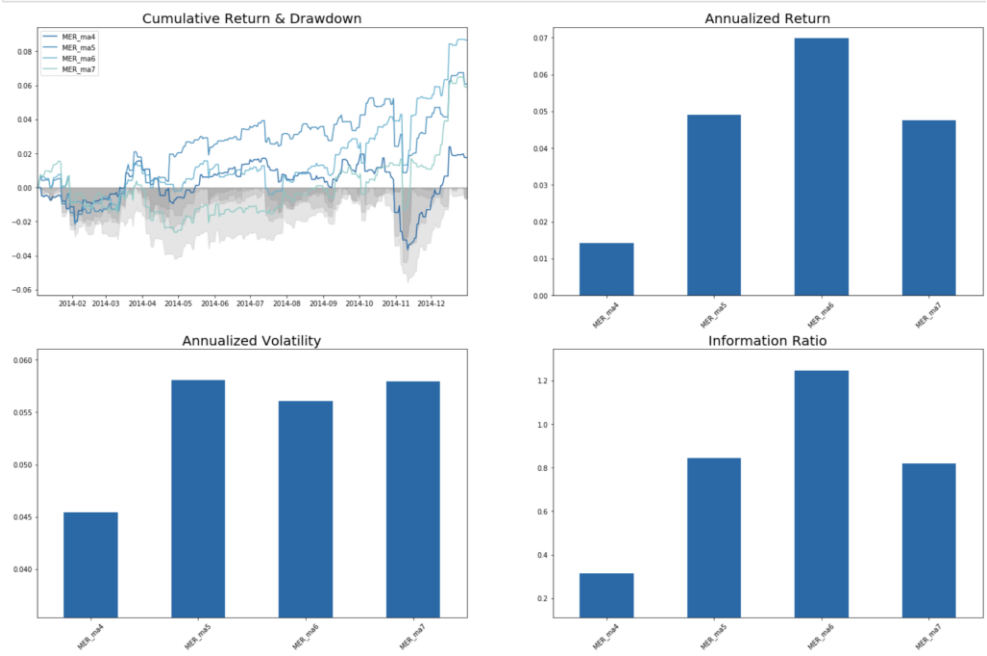




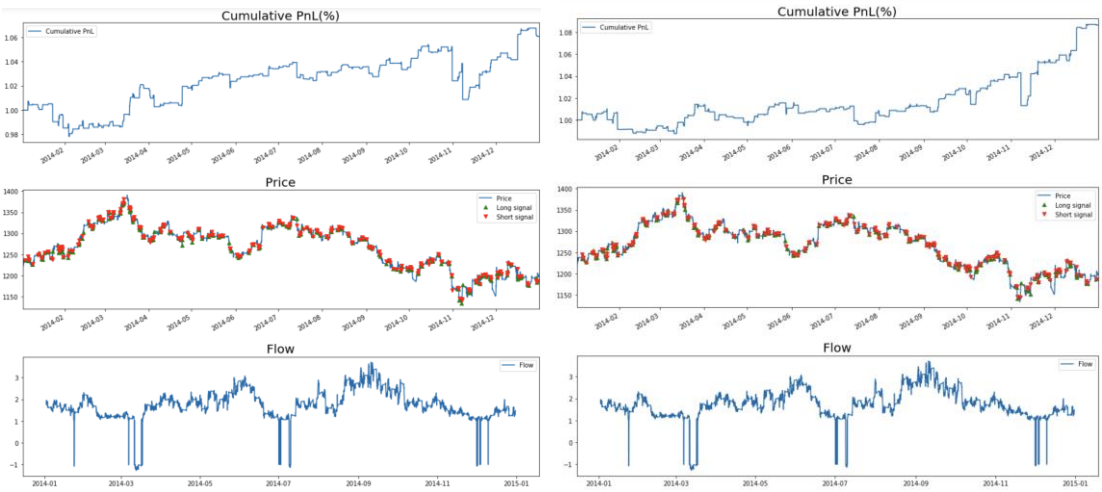
In the end, we lock the parameter N in the range of 4 to 7.

1.5 In-sample Performance of the alpha

Using these parameters for Mean Reversion (MER) strategy in the whole in-sample period, we compute the performance stats for them, as below:

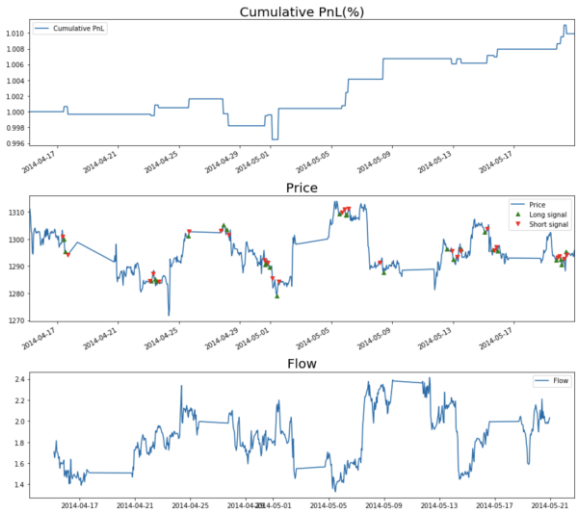


From the graph, where we first consider Information Ratio, we pick up N=5 and N=6 and plot their cumulative return.



When N=5, we may witness a steady profit during 2014-04 to 2014-05, when the price basically in a MER regime, meaning that our strategy successfully captures this pattern. While when N=6, during the same period, the MER strategy gets a loss, which means N=5 is better since it meets our goal.

We can also check at that special one-month period (during 2014-04 to 2014-05), to see the performance of MER5(meaning MER strategy with N=5):

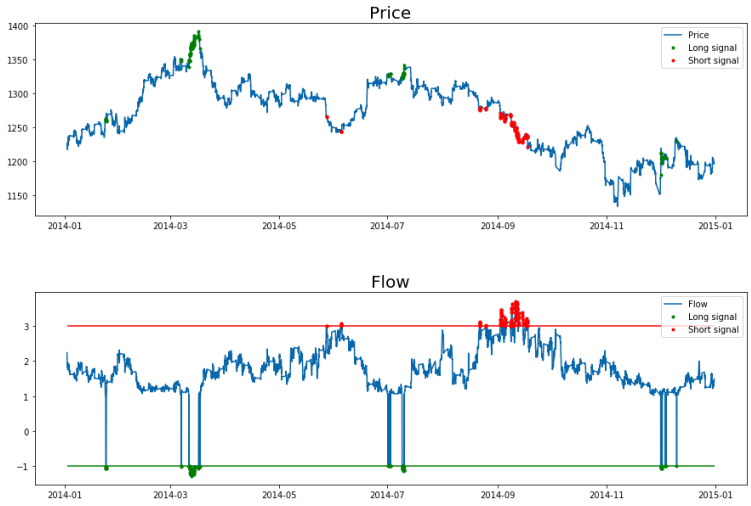


And it actually attains an IR of 5.35 during this specific one-month period, the signals generated from Bollinger bands precisely capture the pattern of mean reversion.

2. Alpha for the Break-Out State

2.1 Identification for the Break-Out state

The break-out state occurs when there exists imbalance in the volume traded, so the price would move out of the past period ranges. Hence, we would use the imbalance between the traded volume to identify the break-out state. Specifically, the flow measures the relative strength between the short and long position. If the flow is not equal to one, it means the long and short positions are imbalanced. Therefore, this could be a signal for the start of the break-out state.

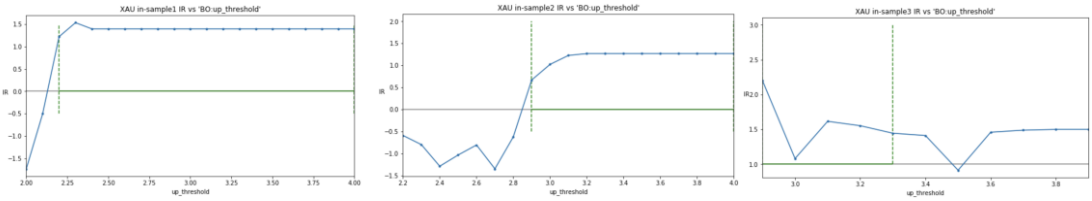


2.2 Strategy for the Break-Out state

Generally, we find that the trend of price vs flow simply goes in opposite direction. When the flow is much greater than one, the price tends to decrease; while the flow is much smaller than negative one, the price would raise significantly. After statistically analysis, we create a certain threshold for the flow to generate our signal for the break-out state. If the position is smaller than -1, the short position is greater than the long position, we would long the asset, so we named it to be the “long threshold”; while if the position is greater than a “short threshold”, when the long position is pretty greater than the short position, we would then short the asset.

2.3 Test for a good parameter for short threshold H in in-sample period

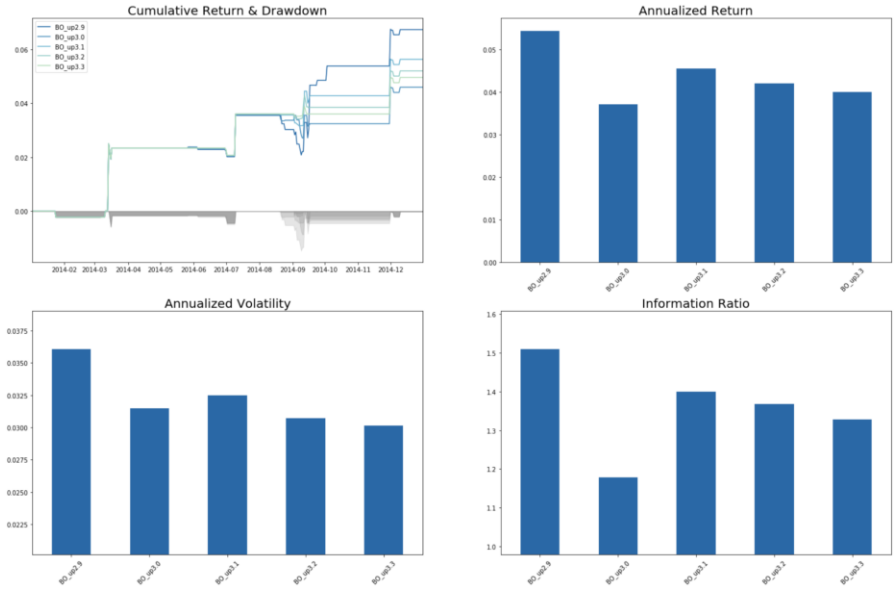
Again, we follow what we do in finding moving average parameter N in MER. The result in in-sample1 (first 1/3 training data), in-sample2(second 1/3 training data), in-sample3(the last 1/3 training data) are shown below:



Eventually we find our short threshold within 2.9 and 3.3.

2.4 In-sample Performance of the alpha

When using Break Out (BO) strategy only, with the H to be 2.9, 3.0, 3.1, 3.2 and 3.3, the in-sample period performance stats are shown below:



Which give us a sense that the BO performs fairly well, especially when H=2.9. That may because of the position data is very helpful for identifying a momentum-based trend in the market, it generates much fewer signals than MER strategies but BO usually get a better precision and earn much for one trade, since it may stay in the market for the entire trend.

3. Combine the Two Alphas

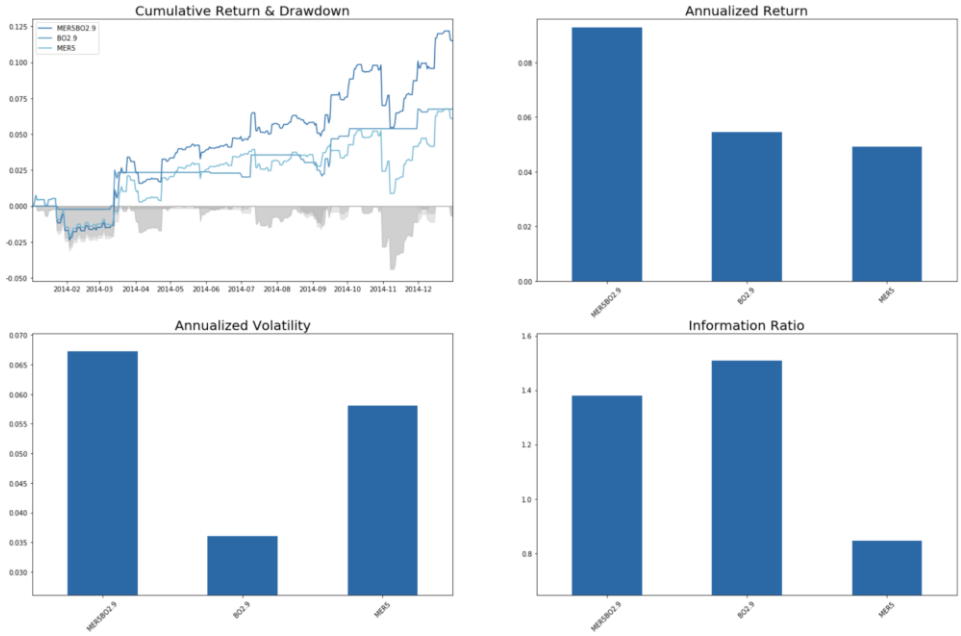
3.1 Priority of the two alphas

The alpha of the Break-Out captures most of the break-out state, when the price would move out of the past period ranges and to a new level. Often, this would go far beyond the level of the mean-reversion state. Therefore, the break-out alpha is prior to the break-out alpha. In detail, if the break-out alpha fires, it would override the signal of the mean-reversion alpha; if the break-out alpha does not signal, the mean-reversion alpha would dominate our trading.

3.2 The Strategy for the XAU_USD

3.2.1 In-sample Selected Strategies for the XAU_USD

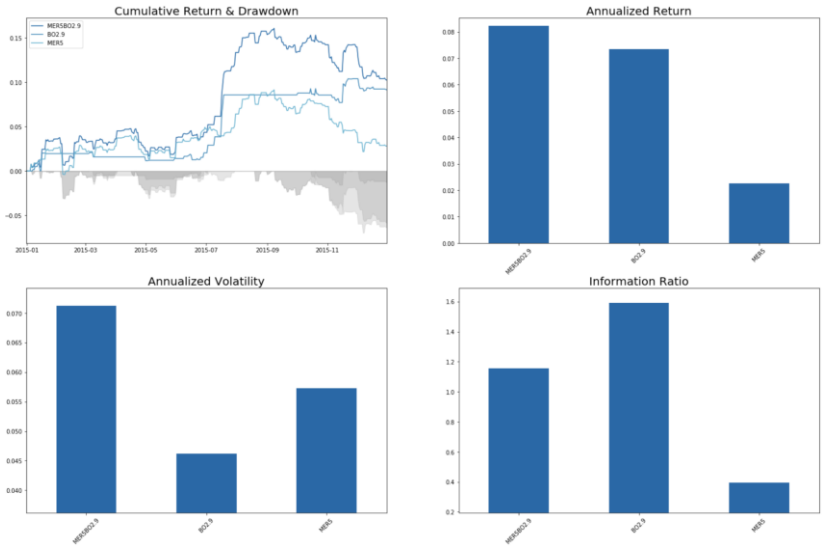
As mentioned above, the window size 5 is the best for the moving average strategy (MER) while the performance of upper threshold of 2.9 is the best for the break-out strategy (BO). Therefore, when combining the two strategies, we set the upper threshold for the break-out strategy to 2.9 with the moving average window fixed to 5. The in-sample result is visualized in the figure below.

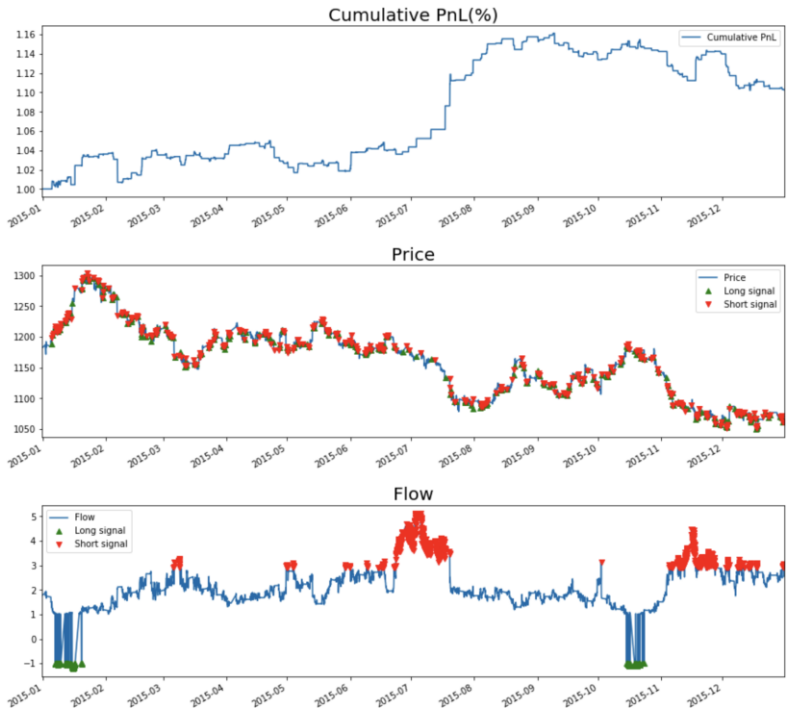


We found that the combination of mean-reversion and break-out alpha and the single break-out alpha have a relatively higher information ratio, the combination of mean-reversion and break-out alpha have a higher annualized return. Hence, the final parameters for our strategy for the XAU_USD is moving average window of 5 and upper threshold of break-out to 2.9.

3.2.2 Out-of-sample Performance of the Strategies for the XAU_USD

After we find our best parameters through the training set, we put our strategies, including MER-BO, BO-only, MER-only, into the out-of-sample test set. The result shows that the information ratio for MER-BO is 1.13, which is satisfactory since our strategy is a high-frequency strategy. Also, the cumulative PnL, the signal of the long and short are plotted in the graph below.





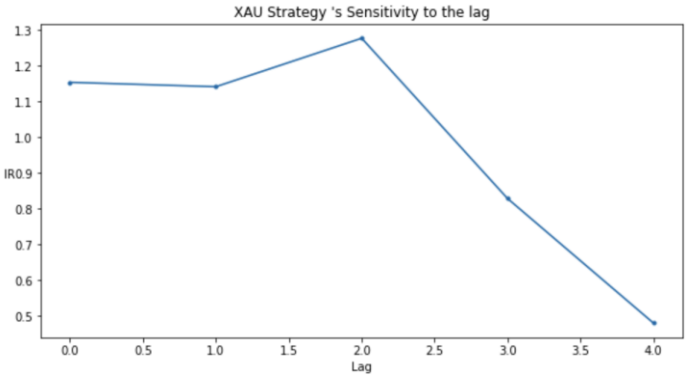
Below are some out of sample trade stats for the MER-BO strategy on XAU_USD.

	Number of Trades	Winner ratio	Winner Mean	Winner Median	Loser ratio	Loser Mean	Loser Median	Maximum DrawDown	Information Ratio
result	352.0	0.5597	0.0024	0.0012	0.4659	-0.0023	-0.0012	-0.0583	1.1545

We may see a winner ratio to be greater than 50% and those winner trades also enjoy a larger mean pnl per trade. Thus, leading to the profitability.

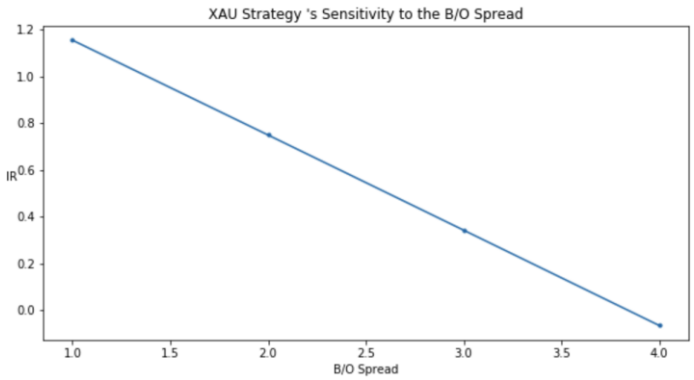
3.2.3 XAU Strategy 's Sensitivity to the lag

Then we did a sensitivity test of the performance of the XAU_USD MER-BO strategy to the lag of execution. As the lag goes longer, the information ratio usually gets smaller, which follows our expectation.



3.2.4 XAU Strategy 's Sensitivity to the Bid/Offer Spread.

Then we did a sensitivity test of the performance of the XAU_USD strategy to the Bid/Offer Spread. As the Bid/Offer Spread goes larger, the information ratio gets smaller, in a linear sense.

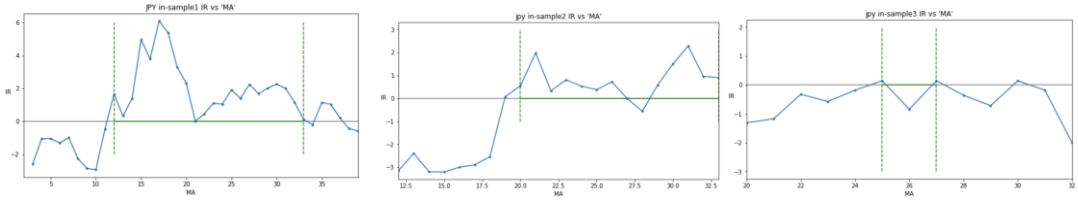


3.3 The Strategy for USD_JPY

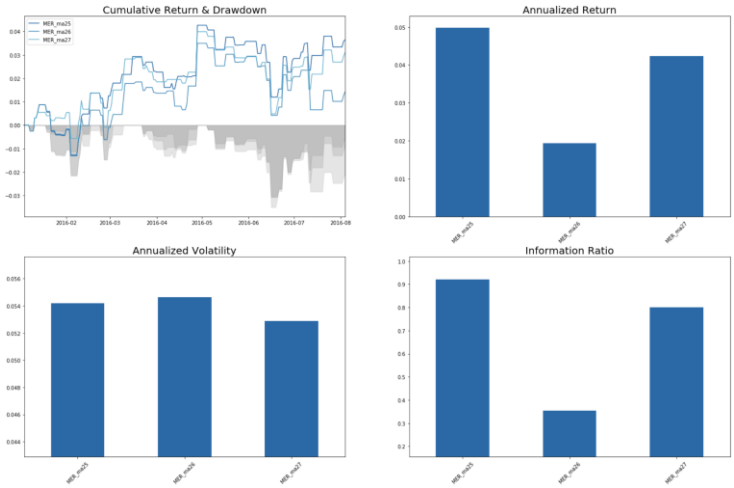
3.3.1 Parameter of the Strategy for the USD_JPY

After finding the best parameters for the XAU_USD strategy, we replicated the same process on the USD_JPY. We first divide the whole data into training set and test set. The training set is the first half data, and the test set is the second half data.

Then, we divide the training data(in-sample) into 3 subsets. Using the each one to run a grid search for N. The performance is judged by Information Ratio, which is the stat considered first in our research. Below shows the in-sample grid search of moving average N:

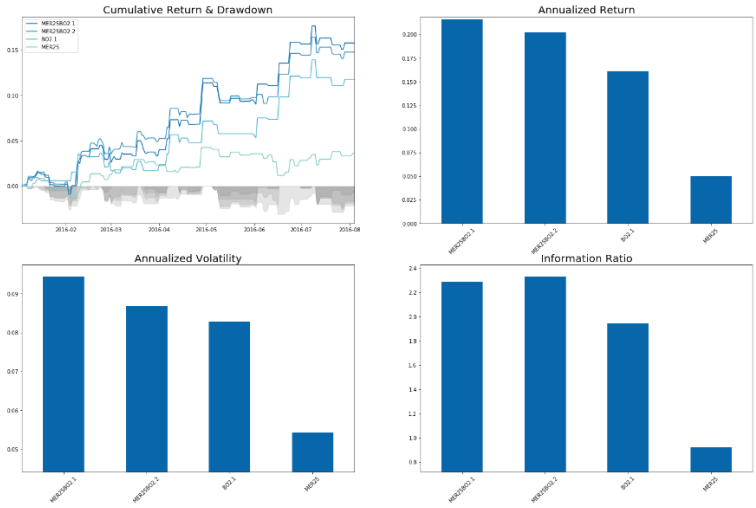


We find the best moving average window range for the USD_JPY is from 25 to 27. Below is the performance of each moving window size in the whole in-sample period.



Based on the information ratio we find the 25 moving window size is the best. And the same logic for the upper threshold for the break out, we find the best is from 2.1.

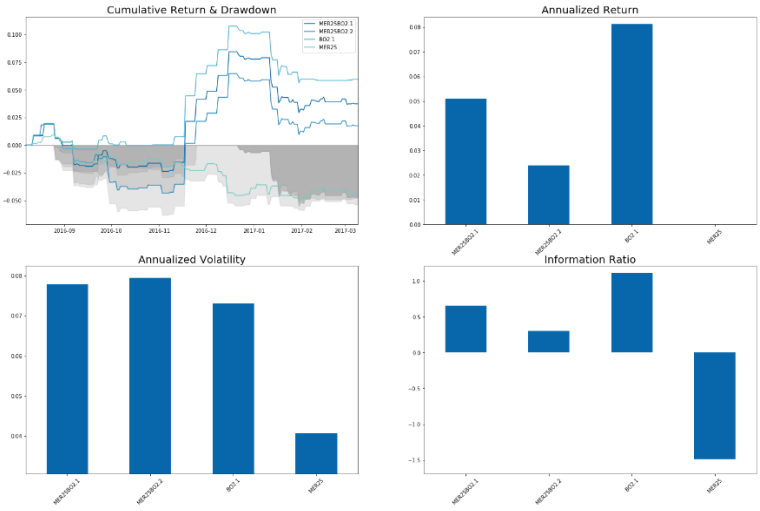
Then we compare MER-BO, BO-only, MER-only strategies with selected parameters during the whole in-sample period, the result is shown below:



Where MER-BO and BO-only perform very well during in-sample period. They all enjoy the IR greater than 2.

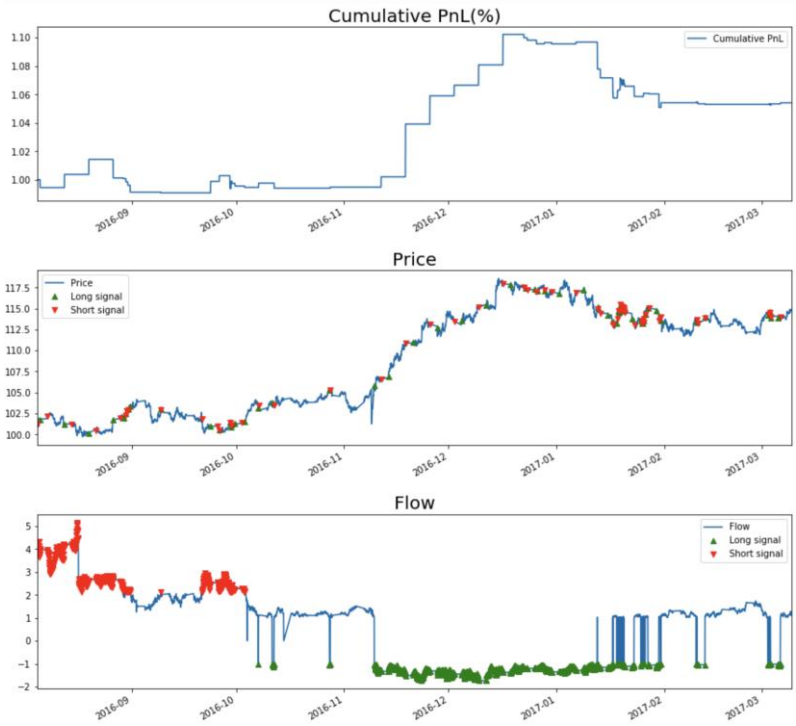
3.3.2 Out-of-sample Performance of the Strategies for the USD_JPY

Then we put these strategies into out-sample data. Below shows the comparison of MER-BO, BO-only, MER-only during the out-sample period.



Here we may see MER-only get a huge loss during out-sample period, which also affects the performance of MER-BO. As a matter of fact, all strategies perform badly during the out-sample period, compared to the performance during in-sample period. In this sense, BO-only gets the best performance for USD_JPY.

The information ratio for BO-only is 1.0. The cumulative PnL for this strategy, and signal of the long and short are plotted in the graph below.



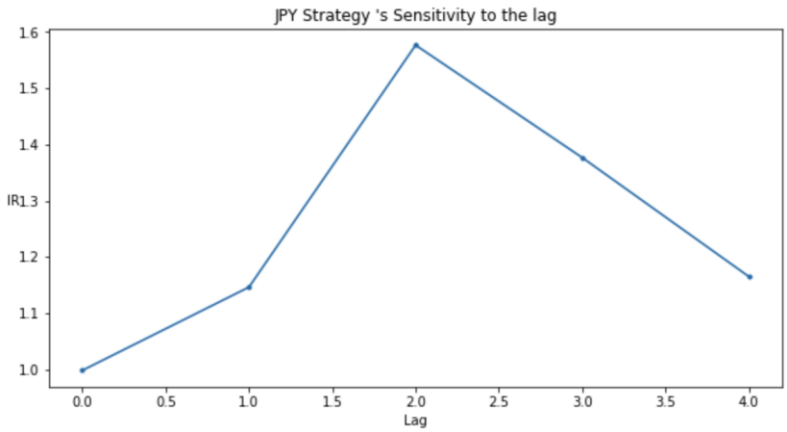
Below are some out of sample trade stats for the strategy on BO-only USD_JPY.

	Number of Trades	Winner ratio	Winner Mean	Winner Median	Loser ratio	Loser Mean	Loser Median	Maximum DrawDown	Information Ratio
result	61.0	0.459	0.0063	0.0035	0.541	-0.0037	-0.0019	-0.0514	0.9982

We may see that the # of trades is small due to it is the BO-only strategy, which generates less signals than MER. Also, the Winner ratio is less than Loser Ratio, while the Winner mean is higher than that of Loser's, meaning that although BO-only strategy generates wrong signals more often but since when it is on the right signals, it can earn more, so the IR is still positive.

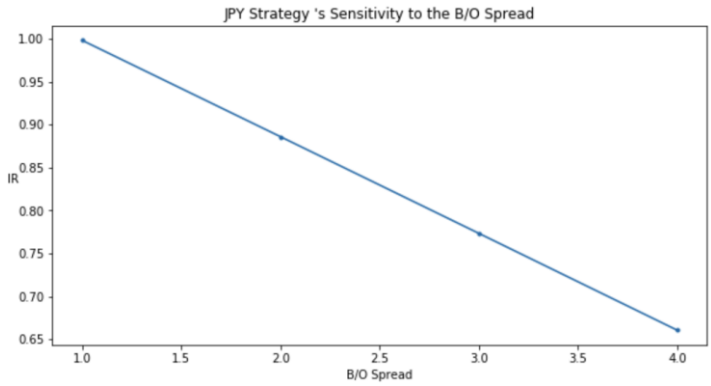
3.3.3 JPY BO-only Strategy 's Sensitivity to the lag

Then we did a sensitivity test of the performance of the USD_JPY BO-only strategy to the lag of execution. As the lag goes longer, surprisingly, we may see an increase of IR when lag goes from 0 to 2, after that the performance becomes to fall. Perhaps, this is because the Break-out state get some latency itself. More specifically, we use Flow data to identify Break-out state and trade after this identification, but the price level may tend to enter its trend several hours later than the time when the Flow data goes to extreme, imbalanced number. And a latency of around 2 hours can capture this structure well during the out-sample period.



3.3.4 JPY BO-only Strategy 's Sensitivity to the Bid/Offer Spread.

Then we did a sensitivity test of the performance of the USD_JPY BO-only strategy to the Bid/Offer Spread. As the Bid/Offer Spread goes larger, the information ratio gets smaller in a linear sense as well.



4. Conclusion

In conclusion, within all our trading strategies (MER-BO, BO-only, MER-only), some of them are pretty effective in out-sample period. For example, the information ratio for the XAU_USD MER-BO strategy in the out-of-sample data is about 1.13, and the information ratio for the USD_JPY BO-only strategy in the out-of-sample is about 1.0. They are all acceptable. What is more, the logic of them is to identify the market regime and traded based on the specific market state. We also consider some real-life detail when we construct these strategies, for instance, we go flat when the market in about to close within one hour and also do not generate further signals during these closing hours.