MODULE 4 UNIT 3
IDE Activity Video 1 Part 2 Transcript

Module 4 Unit 3 IDE Activity Video 1 Part 2 Transcript

STEFAN ZOHREN: So here we follow the implementation which is outlined in detail in this article by Baz from 2015, where we take the MACD signal.

The time series momentum strategy introduced above is a simple trading rule. Since the inception of time series momentum, more sophisticated methods have been proposed. Next, you will consider the work introduced in Baz et al. 2015, which focuses on the volatility-normalised moving average convergence divergence (MACD) indicators.

The trend estimator,
Y_t, and position sizing,
X_t, are obtained through the following steps:

1. Compute the MACD signal as the difference between a short (S) and long (L) EWM average (EWMA) of the price p_t (here (S, L) could be (32, 96) days:
$$MACD(t,S,L) = EWMA(p_t,S) - EWMA(p_t,L)$$
2. Normalise the MACD signal with the EWM standard deviation (EWMSTD) of price of span 63 days.
$$q_t = \frac{MACD(t,S,L)}{EWMSTD(p_t,63)}$$
3. Standardise the signal with an EWM standard deviation of itself of span 252 days.
$$Y_t = \frac{q_t}{EWMSTD(q_t,252)}$$
4. Finally we transform the signal by a squashing/response function
$$X_t = \phi(Y_t), \quad \text{where} \quad \phi(y) = \frac{y\exp\left(\frac{-y^2}{4}\right)}{0.89}.$$
Comments on the response function and its meaning are expanded on below.

Multiple signals with different timescales can also be averaged to give a final position:
$$\tilde{Y_t} = \frac{1}{3}\sum_{k=1}^{3} Y_t(S_k,L_k).$$

So, the MACD signal is effectively just a difference between a short and long EWMA of the price, and if you just process those, you can see that that actually produces some form of momentum. But it's a smoother signal than the one before where we always went up one unit and down one unit. Here, we have more of a continuous signal which, kind of, can have a continuous range. And to, kind of, get a better range, we actually standardise the signal as you see here, so it's nicely waning basically between plus and minus 1 standard deviations for most of the time.

And now, additionally, we perform this squashing function to the signal, and we will see later, I have a plot of this squashing function later below we can see what it does. Now, for this MACD signal, we always have two horizons: we have a short horizon and the long horizon. And we look here at different pairs, so we have the longest one which is, kind of, the long horizon's 96 with 32 days, then we have one where we have 48 and 16 days, and we have a shorter one we have 24 and 8 days.

The time series momentum strategy introduced above is a simple trading rule. Since the inception of time series momentum, more sophisticated methods have been proposed. Next, you will consider the work introduced in Baz et al. 2015, which focuses on the volatility-normalised moving average convergence divergence (MACD) indicators.

The trend estimator,
Y_t, and position sizing,
X_t, are obtained through the following steps:

1. Compute the MACD signal as the difference between a short (S) and long (L) EWM average (EWMA) of the price p_t (here (S, L) could be (32, 96) days:
$$MACD(t,S,L) = EWMA(p_t,S) - EWMA(p_t,L)$$
2. Normalise the MACD signal with the EWM standard deviation (EWMSTD) of price of span 63 days.
$$q_t = \frac{MACD(t,S,L)}{EWMSTD(p_t,63)}$$
3. Standardise the signal with an EWM standard deviation of itself of span 252 days.
$$Y_t = \frac{q_t}{EWMSTD(q_t,252)}$$
4. Finally we transform the signal by a squashing/response function
$$X_t = \phi(Y_t), \quad \text{where} \quad \phi(y) = \frac{y\exp\left(\frac{-y^2}{4}\right)}{0.89}.$$
Comments on the response function and its meaning are expanded on below.

Multiple signals with different timescales can also be averaged to give a final position:
$$\tilde{Y_t} = \frac{1}{3}\sum_{k=1}^{3} Y_t(S_k,L_k).$$

Note that the timescale S (or L), as defined in the above paper, corresponds to `1/alpha` in terms of the smoothing factor `alpha` used in the `pandas.ewm` function. Alternatively, you could also use `span=2*S-1`

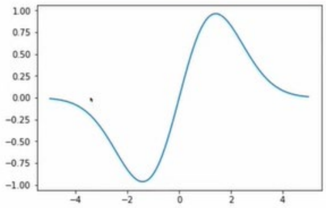
So, in this case, we take all these combinations, we compute the MACD signals, and we now blend all these three signals together.

So here we see a plot of the squashing function.

```
        for short_window, long_window in trend_combinations:

            indiv_signal = self.compute_indiv_signal(prices, short_window, long_window)

            if signal_df is None:
                signal_df = phi(indiv_signal)
            else:
                signal_df += phi(indiv_signal)

        return signal_df / len(trend_combinations)
```

Next, look closer at the response function $\phi(y)$ defined above.

```
In [25]: x = np.linspace(-5,5, 100)
         plt.plot(x,phi(x));
```



You can see that positions are increased until $|Y_t| = \sqrt{2}$, before decreasing back to zero for larger moves. The idea is that you are less certain of larger moves

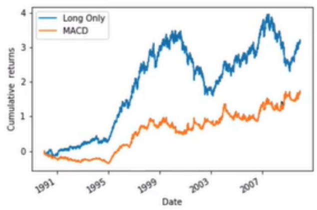
So interestingly, I said to you, because the signal is standardised, most of its values actually ranged between minus 1 and plus 1. Now this squashing function’s actually linear—mostly linear in this regime so we have a nice response to the signal. But what happens is, if we go over the kind of, standard deviation, the squashing function actually goes down again. So, the idea is here that if we have a very large move of the signal, most likely we have seen such large moves only few times in the past, so we aren't really certain about those, so that's why we don't raise them again.

So, this is an interesting way of actually performing a transformation of the signal which takes the effect that in a kind of range of wisdom values of the signals, you want to have a linear response, but if the signal becomes very large, we want to suppress it again because we are very unsure about those values when the signal becomes very extreme. So, we don't want it to take very large bets in those cases. We, again, try to take uncertainty into account here, even though it's done in a very heuristic way.

So, in the following, we now see a plot where we look at the performance of the long-only, and in comparison, we see the MACD strategy which performs less good.

Next, calculate the (volatility-scaled) returns using the combined MACD strategy for different lookback filters and compare these strategies to the long-only strategy.

In [26]:
```
# Calculate returns using MACD
captured_returns_volscaled_macd = (
    MACDStrategy().get_signal(data["srs"])*data["scaled_next_day_returns"]
)["1990-01-01":]

# Plot cumulative returns of MACD vs. Long only
plot_captured_returns(captured_returns_volscaled_lo, plot_with_equal_vol = VOL_TARGET)
plot_captured_returns(captured_returns_volscaled_macd, plot_with_equal_vol = VOL_TARGET)
plt.legend(["Long Only", "MACD"]);
```



Using a combination of signals and averaging them seems to give less than satisfactory results. Consequently, you can study and plot the performances of the filters listed in MACD_TREND_COMBINATIONS instead.

Now, this we will see in the following, has to do with the horizons. So, this case, we had different pairs, different horizon of the strategy, and we ensembled them. And we saw for the standard-time series that the shorter horizons, as well as the combination of the two, also underperformed. So, it is most likely that for the MACD strategy, the longest sector which we gave will also have the best performance.

So, let's just look at this in the following exercise where you were just asked to look at the various MACD trend combinations, and look at them all independently and see how the performance is offset.

Exercise: Calculate the returns using MACD and the different filters within the list MACD_TREND_COMBINATIONS.

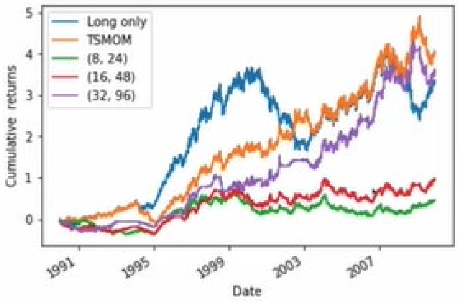
Once calculated, plot the captured returns, rescaling the time series to have 15% volume in order to compare between them.

In [ ]:
```
### Enter code here

# Plot cumulative returns of Long only
plot_captured_returns(captured_returns_volscaled_lo)
# Print performance metrics for long only strategy
print("\nLong Only:")
calculate_statistics(captured_returns_volscaled_lo)

# Plot cumulative returns of TSMOM with lookback of 252
plot_captured_returns(captured_returns_volscaled_tsmom)
print("\nTSMOM: ")
calculate_statistics(captured_returns_volscaled_tsmom)

legends = ["Long only", "TSMOM"]
# Calculate returns using MACD
for windows in MACD_TREND_COMBINATIONS:
    captured_returns_volscaled_macd = (
        MACDStrategy([windows]).get_signal(data["srs"])*data["scaled_next_day_returns"])["1990-01-01":]
    legends.append(str(windows))
    print("\n(S,L) = " + str(windows) + ":")
    print("vol: ")
    plot_captured_returns(captured_returns_volscaled_macd, plot_with_equal_vol = VOL_TARGET)
    calculate_statistics(captured_returns_volscaled_macd)
plt.legend(legends);

###
```

So here, I filled in the code for that. You can go through it in detail in your own time, but by now I think it should hopefully be straightforward.

We can see it's easily seeing the plot here. We can see all the strategies together. We can see the three different MACD combinations together with time-series momentum, and long-only.

```
Sortino Ratio = 0.39
Calmar Ratio = 0.39
Percentage of positive returns = 51.58%
Profit/Loss ratio = 0.989

(S,L) = (32, 96):
vol:
Performance Metrics:
Annualised Returns = 5.83%
Annualised Volatility = 10.04%
Downside Deviation = 7.90%
Maximum Drawdown = 8.39%
Sharpe Ratio = 0.58
Sortino Ratio = 0.74
Calmar Ratio = 0.69
Percentage of positive returns = 52.69%
Profit/Loss ratio = 1.003
```



And here we can see that actually, the longest combination for the MACD performs best and performs quite similar to the time-series momentum strategies. So, we see that MACD and time-series momentum are, kind of, similar in performance here.

So now in the very last bit of the exercise, I wanted to talk a little bit about the use case of momentum for portfolio construction. So, in practice, often some sovereign wealth fund, for example, might employ CTA strategies such as those momentum strategies as a diversifier for their portfolio.

Momentum strategies, as offered by commodity trading advisors (CTAs), have been popular diversifiers in portfolios of, for example, pension funds. Even if returns are small, but the strategy has low correlation with the long-only strategy, you can build portfolios which can outperform long-only strategies in risk-adjusted terms.

Exercise: Portfolio construction

Above, you defined the volatility-scaled long-only and time series momentum strategy and stored the return series in the variables `captured_returns_volscaled_l` and `captured_returns_volscaled_tsmom`, respectively. For this exercise:

a) Compute the correlations between the two return series.

b) Build a portfolio which is equally weighted between the long-only and time series momentum strategy. This makes sense given that the strategies have similar volatility (both target 15%) and similar returns. Alternatively, you can also contsruct the optimial portfolio in a Markowitz sense. Plot the returns of the portfolio and compare it to the individual strategies.

```
In [ ]: ### Part A
        ### Enter code here
```

So, they might be—have a look at long-only investments, and they can do that cheaply themselves by buying ETS, for example. And now they want to diversify their portfolio and we will see that actually, time-series momentum strategies can be quite good because sometimes they are anti-correlated, and we will see that later in the plot as well that during some crisis, like in 2008, times-series momentum was actually anti-correlated. So, while equities markets for example were going down, you will see that the momentum strategy was actually going up. So, if you had both of them in the portfolio, you can actually get better risk adjusted returns.

That's very similar to, for example, a 60/40 portfolio where you have equity in bonds, but here it is done through momentum strategy rather than having bonds, for example. So here in the last exercise, you are asked to just have a look what happens if I mix the long-only strategy with the time-series momentum. And since we only have two instruments here and they have roughly the same volatility because we target 50% for both strategies, we can just equally rate them and blend them together.

Exercise: Portfolio construction

Above, you defined the volatility-scaled long-only and time series momentum strategy and stored the return series in the variables `captured_returns_volscaled_l` and `captured_returns_volscaled_tsmom`, respectively. For this exercise:

a) Compute the correlations between the two return series.

b) Build a portfolio which is equally weighted between the long-only and time series momentum strategy. This makes sense given that the strategies have similar volatility (both target 15%) and similar returns. Alternatively, you can also contsruct the optimial portfolio in a Markowitz sense. Plot the returns of the portfolio and compare it to the individual strategies.

In [ ]:
```
### Part A
### Enter code here

captured_returns_volscaled_lo.corr(captured_returns_volscaled_tsmom)

# We observe around 50% correlation which is not too bad, but ideally could be lower.
# Better results can be obtained on average when looking at larger universes of financial instruments.

###
```

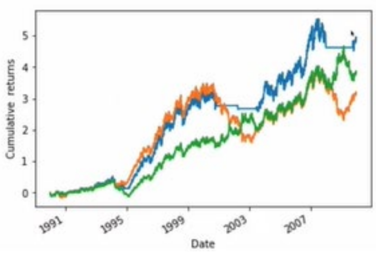
In [ ]:
```
### Part B
### Enter code here
```

If we had more strategies, we had to do something more sophisticated such as a mean variance portfolio, but in this case with two, it's a rather simple exercise.

Now below you can see the plot and you can see that actually, the combination of long-only and time-series momentum performs the best. And the nice thing is, as you can see, is for example, in the kind of .com bubble, after it went down for one year, the time-series momentum strategy actually picked up the downward trend and went short.

```
# We observe overall better risk adjusted returns.
# Interestingly, the TSMOM strategy is anti-correlated during the aftermath of the dot com and subprime crises.
# Thus the portfolio avoids some of the drawdowns the long only strategy sees during those periods.

###
```

```
Performance Metrics:
Annualised Returns = 8.87%
Annualised Volatility = 13.32%
Downside Deviation = 11.05%
Maximum Drawdown = 10.96%
Sharpe Ratio = 0.67
Sortino Ratio = 0.80
Calmar Ratio = 0.81
Percentage of positive returns = 39.90%
Profit/Loss ratio = 0.978
```



While the long-only strategy kept on losing, the time-series momentum strategy kept on gaining, and we actually see a quite flat performance here between 2001 and 2003, which helped to kind of, cut off a little bit of the drawdowns there. And we see a similar behaviour around 2008.

Overall, this led to a better risk adjusted performance. We have lower volatility; before what was 50%, now we only have 30%, and maximum drawdown is down to around 11%. Overall, sharp ratio went up about 0.67. So, we can really see that combining the long-only strategy with the momentum strategy gives us a better risk adjusted performance. And this is actually how in practice, many big pension funds or sovereign wealth funds use those strategies, those momentum strategies offered by hedge funds and CTAs to diversify their portfolio.

I hope this tutorial gave you a basic idea of how systematic strategies work. We have seen how volatility targeting can help to reduce risk and lead to better risk adjusted performance for simple long-only strategies. And we have seen simple examples of momentum strategies and how combining momentum strategies with long-only can help to diversify the portfolio and lead to better overall results. So, this is also how many, for example, pension funds use those type of strategies which are offered by hedge funds to diversify their portfolio and lead to overall better investment results.

Now, obviously in this tutorial, we only focused on a single instrument for simplicity. If you would like to run those type of strategies in real life settings, you need to have a bigger universe of investment instruments, and you have to run those models over those larger universes to get better statistics. So, this was just an example of how it looks like, and if you would like to dig a little bit deeper, I recommend you to read some of the references which we mentioned at the end of the tutorial.

SPEAKER: If you would like to review any of these sections, please click on the relevant button.