MODULE 1 UNIT 2
Video 1 Transcript

Module 1 Unit 2 Video 1 Transcript

NIR VULKAN: We're now going to watch Anthony Ledford, the chief scientist of Man AHL, talking about what patterns look like in real financial data. This is a fascinating presentation, which Anthony often does for investors in AHL, and I've seen that before. We have included a short version here for you to watch, but there is also a longer, more technical version that you can follow by clicking on the appropriate link.

Watch the short version

ANTHONY LEDFORD: My name's Anthony Ledford, I'm the Chief Scientist at Man AHL. Previous to working at AHL, I was a university lecturer in mathematics, and statistics. So, I wanted to show you, today, how you go from observed data, through to actually constructing a signal, and the thing that you really should have in your mind's eye, is that the signal is actually something that's very very weak.

I've taken a hundred different commodities, and futures instruments, and plotted those. So, when you plot prices of markets, you see them on wildly different scales. Some of them are very close to zero, some of them may go up into prices of tens of thousands. So, the first thing to do is to work out how you're going to transform those very very different markets onto a common scale.

Price differences is a very natural thing to start looking at, and that's what I'm going to show on this first plot. It's the price differences from those hundred different markets, and just to make this clearer, each of the different colours you can see corresponds to a different market.

Now, what we've done by looking at price differences if we plot them through time, is actually make them have a fairly stable level. We've stabilised one part of the data, but it's very obvious that we haven't really done anything about the different scales of these underlying data. That's the next thing we're going to address, is how to deal with the very different scales.

So, the next plot shows the volatility of the one hundred different price series I've put together. We divide one by the other, and you end up then with looking at the price changes normalised by the volatility for that market, at that particular time. And, it's at that point that things start to look a lot more stable. You've localised the mean so that it has a level stable around zero, and, regardless of which market is, they're on a sort of common scale of uncertainty around that level of zero. That's the scale where it starts to make sense to do some data modelling, but it's not until you've got it on to that common scale that you can even start fitting models. This is where we're going to focus on: these price changes normalised by what I've called the "predicted vol" – that is what we're going to use in our construction of trading signals.

A technical point here is that I've said I've divided by predicted volatility. That's because I'm dividing today's price move by a volatility measure which is based only on information up until yesterday.

The most commonly used signal for any systematic trading company that has been operating in the commodity trading advisors' space, is the momentum signal, it's really built on this. This is the bread-and-butter signal of momentum, and this has been traded successfully by commodity trading advisors for at least 30 years.

So, to actually set up what we mean by momentum, this is different to the typical momentum effect that's used in cash equity portfolios. I'm explicitly talking about time series momentum. Now, the first thing you need to do is to pin down momentum over some period or scale, and that involves a parameter – I've called it capital "L" here – and it's called, usually termed, the "lookback". Now, if we denote the Volatility Normalised Price Differences, that we calculated in the previous plot, by capital R, then what we're going to do is take weighted combinations of those values.

And, the time series momentum is just a linear sum of weights multiplied by previous values of those volatility normalised returns – the formula's on the sheet there. The weights W i, in the formula, they are chosen so that the momentum has a variance of one. You don't need to do that, but it makes tremendous sense to do that if you want to look at things on a common scale when you come to analyse different lookbacks – you need to have some sort of familiar measure of what you're looking at. If the scale is changing all the time, you can't really learn anything from that, so, I always make the point that, try and standardise the variance so that you can interpret the plots.

Other considerations are important. We haven't mentioned turnover, but that is something that's very important in terms of the transaction costs you'll incur, and so there's a large effort placed on making the choice of the weights in such a way that it minimizes the turnover.

Now, the point about this signal is it's being used very successfully, very broadly, by a large number of companies from a long period of time. That's because it's actually a very strong signal compared to a lot of the more subtle signals that we try and extract and use within systematic trading today. If you bear in mind this is actually one of the strongest signals, what would you expect if we were to plot, for example, the momentum calculated in this way, against the future price movement? That's what we're trying to use it for after all. It's a way of predicting future price movements.

Now, this is the thing I want you to get in your mind's eye: it's an incredibly subtle effect. If it was a very obvious effect, you might expect to see some straight line. It doesn't look like a straight line. In fact, it's very hard to distinguish there's any signal at all.

This is what the plot looks like, it just looks like noise, effectively. Now, if we're going to actually extract what the signal… its value is for predicting future returns, we need to do some work on this data.

Before I move on to show you how to do that, just note the vertical scale. This plot goes from minus 4 to plus 4. That's very important. That's the scale at which the data is operating. That's pretty much the scale at which the noise operates. We're going to extract a signal from this, and you'll see it's tiny compared to that minus 4 to 4 scale.

Now, the way we're going to do this is to take this plot and slice it into vertical strips that have the same number of points in each of those vertical strips, and then we'll take the mean, and the standard deviation, within each of those vertical strips, and plot that. If you do that, you get a plot where it does look like there is some sort of structural predictability from this momentum signal, but the scale of that predictability is tiny compared to the scale of the original plot.

Here you'll see that it goes from minus 0.1 up to 0.1, or thereabouts. So, the original data was on the scale of minus 4 to 4, and the signal is on a scale of minus 0.1 to 0.1. Now, this is not the full data set. This is all of the commodities out of those a hundred markets. But this is the type of effect that we see when we've put the commodities on to the same scale and tried to process the data to actually extract what the signal looks like.

We can do the same in other sectors. It's kind of a feature of markets rather than a feature of, say, stocks, or bonds, or FX. It's very very prevalent, and that's why it's been used so successfully by many firms over the last multiple decades.

So, we've looked at the commodities. If we go back and do the same analysis, the same way of chopping the data, over all of the one hundred markets… again, put all the markets together, you can't really see anything. It just looks like noise. You do our calculation of the means and standard deviations within each vertical strip. And remember, each of those strips has the same number of points in it, by construction – that's the way I've constructed those strips – you again see this effect. It goes from minus 0.5 to plus 0.5 on this plot, but that is the nature of what we're trying to do. And remember, this is one of the strongest signals you will find, and it's barely detectable with the amount of noise that's there.

That's so important, you know, don't expect the answer to leap off the page. It won't, you have to work very hard to see anything in terms of structure, and then you've got to do a lot more work as well to make sure that the thing that you've observed from your data actually persists. Look at it within different five-year regions, or different one- or two-year regions. Those are the necessary steps you need to do in order to understand the structure of these data for building a trading signal.

So far, all we've done is look at just a single, a single momentum signal, but there is far more that you need to consider.

We haven't thought about how do we actually go and turn that effect into something that involves taking a position in a financial market, and that's a major thing you need to think about.

The first thing to think about is, well, how do you respond to that signal? The other thing to think about, is cost reduction. I think about these things as taking an optimal position based on the latest data. Now, as new data comes in, my view of what the optimal position is will change. We've only spoken about a single momentum system so far, but nobody would ever trade a single momentum system on a single market, and expect to make money from that.

You need to think of a whole spectrum, a whole range of different lookbacks – that will give you some sort of diversification with respect to the time scale or lookback parameter. And, nobody would ever trade it on a single market as well. Most large commodity trading advisors have portfolios that involve hundreds of markets.

There is a whole range of other things that you can look at as well, and I've listed some of those here: breakouts, carry, calendar effects, etc. I'll let you look through up. And more recently, people are talking about extracting signals from alternative data, of course.

They're looking at, for example, footfall going through retail parks, or looking at extracting effects from social media.

There's a vast amount of data out there. There's a vast amount of analysis you can do.

The last thing I would say is use your common sense, scrutinise your results. Really ask the question, does this make sense to you? Because there is tremendous risk with all this data, that you can overfit systems, you start believing and finding effects that aren't really there, they're just curiosities of the noise in the data. You have to be incredibly careful with very powerful models, and very broad data sets, that you are actually capturing something that's genuine. There's a paper I've listed there, the full references given, this explains the effect quite nicely. It's called "Dot, dot, dot, and the cross section of expected returns", and it was written by Campbell Harvey, and two co-authors. That raises the question that the many many different effects that have been identified, and supposedly scrutinised, and peer-reviewed, and published in the literature, makes the very valid point: they can't all be real. And I'll leave you to look at that one.

Always scrutinise, never jump to believing. Data mining can be done well, but it can also come back to bite you.

NIR VULKAN: If you want to see a longer, more technical version of Anthony Ledford notes, or if you prefer to go to the conclusion, click on the corresponding button.

Watch the longer version

ANTHONY LEDFORD: My name's Anthony Ledford, I'm the chief scientist at Man AHL. Previous to working at AHL, I was a university lecturer in mathematics, and statistics. So, I wanted to show you, today, how you go from observed data, through to actually constructing a signal, and the thing that you really should have in your mind's eye, is that the signal is actually something that's very very weak. The amount of noise that's there compared to the amount of predictable signal is huge. This is not like a lot of other areas of data analysis, where the noise is actually not capable of dwarfing the signal. Here, you have to work really hard to actually find any signal at all.

So, I'm going to start off by just plotting a series of prices, and I've taken a hundred different commodities and futures instruments, and plotted those. That's the other thing about algorithmic trading: most algorithmic traders try and do things at a scale in terms of breadth. They don't build portfolios that are very concentrated around a few instruments – that's more what discretionary traders do – and they have a very deep understanding of a few particular sectors and individual markets. That's not what algorithmic traders do. Their job is to really produce portfolios that are very broad. That's why I'm doing something here with a hundred different markets.

So, when you plot prices of markets, you see them on wildly different scales. Some of them are very close to zero. Some of them may go up into prices of tens of thousands. Before you can do any sensible modelling on that, you need to get the data onto some sort of common scale, or else you're not going to be able to model anything, you're comparing apples with oranges. So, the first thing to do is to work out how you're going to transform those very very different markets onto a common scale.

Typically what we're interested in doing in building systematic trading systems, is understanding predictability of where prices are going. Are they going up or down? So, price differences is a very natural thing to start looking at, and that's what I'm going to show on this first plot. It's the price differences from those a hundred different markets, and just to make this clearer each of the different colours you can see corresponds to a different market.

Now, what we've done by looking at price differences if we plot them through time, is actually make them have a fairly stable level. So, rather than the price drifting up or drifting down, the price difference tends to be centred around zero. So, we've stabilised one part of the data, but it's very obvious that we haven't really done anything about the different scales of these underlying data.

The prices are still moving at very different scales. That's the next thing we're going to address, is how to deal with the very different scales at which the prices move from market to market, and also, within the same market, sometimes prices are moving only a small amount, and sometimes they're moving a large amount, and this effect is called volatility. And that's the next thing we need, the next ingredient we need to actually do some data modelling here.

So, the next plot shows the volatility of the one hundred different price series I've put together. And, you can see, again, the different lines – if you look at lines of the same colour, they go through periods of low volatility, and some periods of high volatility, and there's a big difference from market to market there.

Once we've got those two components, the price differences and a measure of the volatility, we divide one by the other, and you end up then with looking at the price changes normalised by the volatility for that market at that particular time. And it's at that point that things start to look a lot more stable. You've localised the mean so that it has a level stable around zero, and regardless of which market is, they're on a sort of common scale of uncertainty around that level of zero. That's the scale where it starts to make sense to do some data modelling, but it's not until you've got it on to that common scale that you can even start fitting models. So, this is where we're going to focus on: these price changes normalised by what I've called the "predicted vol" – that is what we're going to use in our construction of trading signals. And typically, that's where most analyses will start.

A technical point here is that I've said I've divided by predicted volatility. That's because I'm dividing today's price move by a volatility measure, which is based only on information up until yesterday. I'm not allowed to have any forward leakage of information here – but that's the scale we're going to work on. Sometimes those are called "volatility normalised returns".

Now, the most commonly used signal for any systematic trading company that has been operating in the commodity trading advisors' space, is the momentum signal, is really built on this. So, that's the simplest thing we're going to look at. There's a whole plethora of different signals that I will mention later on, but just to be absolutely clear what we're doing, this is the bread-and-butter signal of momentum, and this has been traded successfully by commodity trading advisors for at least 30 years. We've been doing it for over 30 years.

So, to actually set up what we mean by momentum, I'm going to be, again, a little bit pedantic, and I'm going to talk about time series momentum here, because if you're an equity trader trading cash equity portfolios, if you talk about momentum in cash equities, it means something completely different, that's a cross-sectional effect. This is different to the typical momentum effect that's used in cash equity portfolios. I'm explicitly talking about time series momentum; and for the technically minded, this is a univariate signal, it is created using the history of that market, and that market alone. It is not cross-sectional signal.

Now, the first thing you need to do is to pin down momentum over some period or scale, and that involves a parameter – I've called it capital "L" here – and it's called, usually termed, the "lookback". Now, if we denote the Volatility Normalised Price Differences, that we calculated in the previous plot, by capital R, then what we're going to do is take weighted combinations of those values.

And the time series momentum is just a linear sum of weights multiplied by previous values of those volatility normalised returns – the formula's on the sheet there. The weights W i, in the formula, they are chosen so that the momentum has a variance of one. You don't need to do that, but it makes tremendous sense to do that if you want to look at things on a common scale when you come to analyse different lookbacks – you need to have some sort of familiar measure of what you're looking at. If the scale is changing all the time, you can't really learn anything from that, so I always make the point that, try and standardise the variance so that you can interpret the plots.

Other considerations are important. We haven't mentioned turnover, but that is something that's very important in terms of the transaction costs you'll incur. And so, there's a large effort placed on making the choice of the weights in such a way that it minimizes the turnover. We'll come on to that more in future. Now, the point about this signal is it's being used very successfully, very broadly, by a large number of companies, for a long period of time. That's because it's actually a very strong signal, compared to a lot of the more subtle signals that we try and extract and use within systematic trading today. If you bear in mind, this is actually one of the strongest signals, what would you expect if we were to plot, for example, the momentum, calculated in this way, against the future price movement? That's what we're trying to use it for after all, it's a way of predicting future price movements.

Now, this is the thing I want you to get in your mind's eye: it's an incredibly subtle effect. If it was a very obvious effect, you might expect to see some straight line. It doesn't look like a straight line. In fact, it's very hard to distinguish there's any signal at all. This is what the plot looks like, it just looks like noise, effectively. Now, if we're going to actually extract what the signal, its value is for predicting future returns, we need to do some work on this data.

Before I move on to show you how to do that, just note the vertical scale. This plot goes from minus 4 to plus 4. That's very important. That's the scale at which the data is operating. That's pretty much the scale at which the noise operates. We're going to extract a signal from this, and you'll see it's tiny compared to that minus 4 to 4 scale.

Now, the way we're going to do this is to take this plot and slice it into vertical strips that have the same number of points in each of those vertical strips, and then we'll take the mean, and the standard deviation within each of those vertical strips, and plot that. If you do that, you get a plot where it does look like there is some sort of structural predictability from this momentum signal, but the scale of that predictability is tiny compared to the scale of the original plot.

Here, you'll see that it goes from minus 0.1 up to 0.1, or thereabouts. So, the original data was on the scale of minus 4 to 4, and the signal is on a scale of minus 0.1 to 0.1.

Now, this is not the full data set. This is all of the commodities out of those a hundred markets, but this is the type of effect that we see when we've put the commodities on to the same scale and tried to process the data to actually extract what the signal looks like. We can do the same in other sectors.

Here, I've plotted the FX data. Again, it just looks like noise to your eye. We have to do some work on this. Notice also, that the vertical scale has gone from minus 4 to 4 again. So, even though this data comes from a completely different sector, this standardisation of the original data by volatility normalising it, gives us something that is on a common yardstick, a common frame of reference – that's really important to understand the structure of the data.

So, if we do the same process again by slicing it into vertical strips and calculating the mean, and the standard deviation within those strips, you see a plot that looks like this. It's subtly different to the one we had for the previous sector, the commodity's sector, but broadly speaking it's doing the same thing. And this is why momentum is very very useful. It's because it's a subtle, but persistent and very widespread effect across the whole range of markets. It's kind of a feature of markets, rather than a feature of, say, stocks, or bonds, or FX. It's very very prevalent, and that's why it's been used so successfully by many firms over the last multiple decades.

So, we've looked at the commodities, we've looked at the FX sector, if we go back and do the same analysis, the same way of chopping the data, over all of the one hundred markets… again, put all the markets together, you can't really see anything. It just looks like noise. You do our calculation of the means and standard deviations within each vertical strip. And remember, each of those strips has the same number of points in it, by construction – that's the way I've constructed those strips – you again see this effect. There is a small effect, and it's tiny compared to the amount of noise there. It goes from minus 0.5 to plus 0.5 on this plot, but that is the nature of what we're trying to do. And remember, this is one of the strongest signals you will find, and it's barely detectable with the amount of noise that's there.

That's so important, you know, don't expect the answer to leap off the page – it won't, you have to work very hard to see anything in terms of structure, and then you've got to do a lot more work as well to make sure that the thing that you've observed from your data, actually persists. Look at it within different five-year regions, or different one- or two-year regions, for example, and see is the effect persistent? Is it diminishing? Is there some sort of change going on? Those are the necessary steps you need to do in order to understand the structure of these data for building a trading signal.

There's more you can do with this data than just look at the conditional means. There's a subtle feature of it as well, if you look at the standard deviation of those blocks within each vertical strip, you see that there is some structural change there. Around the momentum being zero, you actually have a variability that's around one, but as you get into higher values of momentum on the positive and negative side, you start to see that there is a growth, almost like a parabolic effect, in the uncertainty.

Now, this is nothing to do with sample size, remember, because, by construction, I've constructed those vertical strips to have the same number of points in each one. This is actually a subtle feature of the uncertainty on that normalised variance scale, conditioning on the momentum. That's possibly more technical than most of you need to know, but once you've got these two things – the conditional mean and the conditional volatility – you're actually very well-equipped then to start putting that together, in a sensible way, to make an efficient trading system portfolio.

There's a very good paper that was written around the year 2000 by Ferson and Siegel, published in the Journal of Finance – the reference is on the page here – that tells you, once you've got those two components, the conditional mean and the conditional standard deviation for any predictor you may find, how to put that information together in such a way to make an efficient portfolio. What I mean by efficient is the one that will give you the best return for the minimum amount of risk, that recipe is detailed in that paper. I'm not going to say any more about it here, but it's not about momentum, that's about how to use if… any edge you've been able to find, how to use it efficiently to build a portfolio. I'll leave you to look at the details of that paper.

So far all we've done is look at just a single, a single momentum signal, but there is far more that you need to consider if you're going to go from building a single signal before you can actually turn that into a full trading system, and I'm going to run through some of those things here.

The first thing is we've made a signal, we've shown that there's some predictability from that signal to future prices, but we haven't thought about how do we actually go and turn that effect into something that involves taking a position in a financial market, and that's a major thing you need to think about.

The first thing to think about is, well, how do you respond to that signal? Should you just follow it, for example, as a binary suggestion? So, if it's negative you should be short one unit, and if it's positive you should be long one unit. Twenty years ago, a lot of people did that, but that's not, that's not the appropriate level of sophistication to get the most value out of these things. And that leads you into thinking about response functions. Should your response function be linear? Should you cap the linear effective so? Should you actually fit a nonlinear response function? All of these things the things you need to think about.

The other thing is volatility scaling. This is incredibly important, because you don't want your risk to be affected, particularly, by nuisance effects that you have no control over because the market's volatility has changed. We've already seen from the plots that volatility in any particular financial market can, and does, change wildly.

So, if you have a particular position in a market – doesn't matter if it's long or short – if the volatility is low one day, and the volatility is high the next day, you're going to be taking very different amounts of risk, and that is probably not something that's useful to you. So, if you want to neutralise that effect, you need to scale your position inversely by the market's volatility, and that's a very important effect to put into systematic trading systems.

The other thing to think about is cost reduction. Now, a lot of emphasis is placed on this kind of works about making trading systems. I actually don't think that's the right way to think about it at all. I think about these things as taking an optimal position based on the latest data. Now, as new data comes in, my view of what the optimal position is will change, and as it changes from one day to the next, I will have a different view of what the optimal position is. So, if yesterday's optimal position was 100, and today's optimal position is 99, there's some discrepancy there. Do I do that trade? Well, you have to ask yourself the question, is a small adjustment to this position actually useful to you, or is it just noise?

You typically want to filter out the noise trades. You don't want to do them, it's just going to incur transaction costs for you. But, if there's a big change between yesterday's optimal position and today's optimal position, you probably do want to trade, so you need to do some sort of filtering, removing the spurious noise trades. That's why I see trades as being a by-product of maintaining an optimal position, with some sort of cost reduction around it.

We've only spoken about a single momentum system so far, but nobody would ever trade a single momentum system on a single market, and expect to make money from that – it's too subtle in effect. So, you need to think about not just a single lookback, you need to think of a whole spectrum, a whole range of different lookbacks, that will give you some sort of diversification, with respect to the time scale or lookback parameter. And nobody would ever trade it on a single market as well; most large commodity trading advisors have portfolios that involve hundreds of markets. Again, it's about diversification there. So, those two things are incredibly important.

And then there's the capacity. Typically, these things are run as a fund management business, rather than as a prop-trading enterprise, and that means that you have a large amount of capital, and you have to deploy that in such a way across the different systems, the different speeds, and the different transaction costs in those markets. Some markets are very cheap. Some markets are very expensive to trade. Stock markets tend to be cheap, FX rates tend to be cheap in developed markets. Commodities, on the other hand, tend to be quite expensive. So you need to understand the transaction costs, the turnover of your system, and the liquidity of those markets, so that you spread your risk out appropriately across that full range, or else you will end up either taking too much of frictional effect through transaction costs, or you will not get your optimal allocation across the portfolio correct. Those things are all incredibly important.

We've really focused on just the momentum sector so far, but momentum is just one group of predictors. They're very popular. They're very persistent. They've been used for years, but that is just one slice of the pie. There is a whole range of other things that you can look at as well, and I've listed some of those here: breakouts, carry, calendar effects, etc. I'll let you look those up. More recently, people are talking about extracting signals from alternative data, of course. They're looking at, for example, footfall going through retail parks, or looking at extracting effects from social media.

There's a vast amount of data out there. There's a vast amount of analysis you can do.

The last thing I would say is use your common sense, scrutinise your results. Really ask the question, does this make sense to you? Because there is tremendous risk with all this data, that you can overfit systems, you start believing and finding effects that aren't really there, they're just curiosities of the noise in the data. You have to be incredibly careful with very powerful models, and very broad data sets, that you are actually capturing something that's genuine. There's a paper I've listed there, the full references given, this explains the effect quite nicely. It's called "Dot, dot, dot, and the cross section of expected returns", and it was written by Campbell Harvey, and two co-authors. That raises the question that the many many different effects that have been identified, and supposedly scrutinised, and peer-reviewed, and published in the literature, makes the very valid point: they can't all be real. And I'll leave you to look at that one.

Always scrutinise, never jump to believing. Data mining can be done well, but it can also come back to bite you.

NIR VULKAN: Thank you for your insight, Anthony. I'm sure they will be valuable for our participants throughout this programmes, and beyond.

If you'd like to review any of these sections, please click on the relevant button.