ON THE EFFICIENT MARKETS HYPOTHESIS

An economist and his friend are walking down the street when they come across a hundred-dollar bill lying on the ground. The friend bends down to pick it up, but the economist stops him, saying, "***Don't bother – if it were a real hundred-dollar bill, someone would have already picked it up***"

WELCOME TO QUANTITATIVE TRADING STRATEGIES!

- This course will proceed over 10 sessions with 1 session reserved for student presentations of final projects
- In this class, we will survey trading strategies across multiple asset classes that are commonly implemented in industry today.
- Many of these strategies are also described and outlined in academic research papers. We will therefore use many papers from published research to illustrate the strategies' theory
- Most strategies have been replicated in Python, together with a multi-purpose backtesting simulator which has also been written in Python. This will be provided together with each class's slides and backend data to aid strategy replication

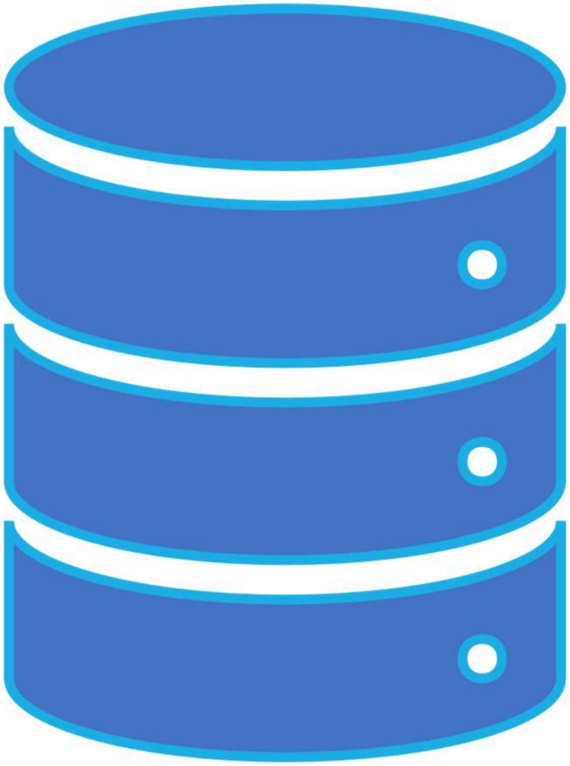
COURSE ADMINISTRATION

Data for selected strategies have already been uploaded to course folders. Please let me know via email after this class if you are unable to access; I will email a link to you directly

Selected strategies will be replicated by me in python code, and then shared with the class. In past iterations of this course, replication was done using third party platforms such as Quantopian, QuantConnect, etc. Intention was to save time having to download and merge data and having code up your own backtesting simulator. However, two disadvantages of this approach are (i) a significant amount of time is needed to 'learn' how to use the third party platform in the first place and (ii) being platform dependent has its own disvantanges – for e.g., Quantopian now does not exist and students need to migrate code. More importantly, **given modern python libraries, merging data and coding up a strategy backtesting simulator is literally a 10 to 20 minutes effort nowadays, and we will have greater control after that (in any case, I will share the backtesting simulator code in class 2).**

For the 2024 version of this class, we will improve on previous approaches by being platform independent. All data will be provided in csv only, not other proprietary format. In addition, sample strategies will be provided in python code that only needs numpy and pandas to run. You **should only need a python interpreter run all the example code I will provide**

<u>**Try to install a Python interpreter by our next class**</u>



COURSE ADMINISTRATION

Refer to the course outline for grading breakdown. Note group project and final exam together make up more than 50% of course assessment

On class participation, you are encouraged to ask questions and provide comments. You are strongly encouraged to do so during regular class time where the rest of your classmates can benefit from the discussion. For offline communication, please use the discussion forum, so that question and answers are in public view.

From class 2 onwards, selected students will be asked to prepare and verbalize 1 to 2 minute summaries of topics that we had described in previous classes. You will be notified in advance if this is something you need to do for a specific class. By end of this course, it is expected that all students will have participated in this activity

COURSE ADMINISTRATION

On class project:

- Try to get formed into teams by around class 3
- For past iterations of this class, average group size has been around 3 to 5
- All group members need to participate in the group project; there are many different activities needed, from planning and literature survey, gathering data, writing code, presentation development, presentation execution and writing of final report

COURSE ADMINISTRATION

Class project (continued):

- The project should implement a quantitative trading strategy.
- It can be over any asset class you chose, and also in any trading style
- Broadly speaking, we will look in equal parts at (i) fundamental economic justification, (ii) implementation of actual strategy, (iii) quality of the backtest and also how realistic it is, and lastly (iv) discussion of performance metrics
- There should be sufficient code and sample data provided in the course of the next few weeks so that you can get a good start by adapting this
- We will discuss the project more in class 2 or 3

Final exam:

- This will be a mix of structured and open ended problem solving, with some MCQs added to test basic knowledge


Memory Lane
A trip down
memory lane

MOTIVATION FOR QUANTITATIVE TRADING

Quantitative trading strategies are often systematic, which **implies performance can be expected to be (more) repeatable**

Has performance (return-drawdown characteristics) that are quantifiable and known in advance via backtesting. This allows us to **calibrate risk exposure (e.g. leverage) better in order to maximize returns**

May be based on **a greater number of instruments** relative to discretionary portfolios

BETA TO 'MARKET PORTFOLIO' DOMINATED OUR UNDERSTANDING OF STOCK RETURNS IN THE 1960S

Modern portfolio theory in the 1960s was focused on 'capital asset pricing model (CAPM)', centered around cross sectional relationship between individual security returns and market returns:

$$E(R_i) = R_f + \beta_i \left[E(R_M) - R_f \right],$$

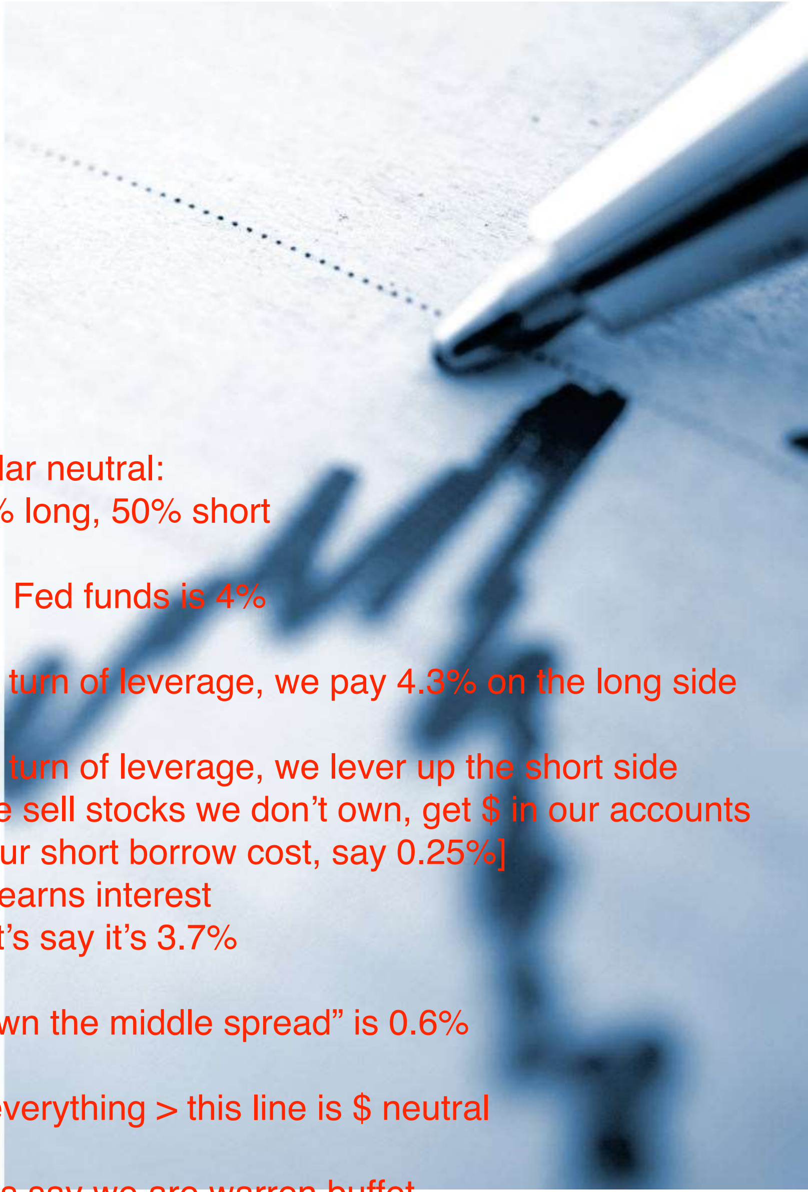
"RULES BASED TRADING" SAW ITS START WITH PUBLICATION OF 'MARKET ANOMALIES' AROUND LATE 1970S, AND THEN EXPANDED INTO FACTORS

Cracks in the CAPM appeared in the late 1970s and early 1980s.

Basu (1977) reported **portfolios of low P/E stocks earned (on average) around 7% more annual than portfolios of high P/E stocks**

Banz (1981) and Reinganum (1981) **reported differences in average returns of portfolios grouped by market cap, unrelated to estimated betas**

Keim (1983) found that the **size effect had a strong seasonal component concentrated in January ("January effect")**



Dollar neutral:
50% long, 50% short

e.g. Fed funds is 4%

Per turn of leverage, we pay 4.3% on the long side

Per turn of leverage, we lever up the short side
* we sell stocks we don't own, get $ in our accounts [incur short borrow cost, say 0.25%]
* $ earns interest
* let's say it's 3.7%

"down the middle spread" is 0.6%

__everything > this line is $ neutral

Let's say we are warren buffet

* long only
* every turn of leverage, we pay 4% (but it varies with fed funds rate)

THE LATE 1970S AND EARLY 80S SAW AN INCREASING NUMBER OF ANOMALIES PUBLISHED IN ACADEMIA

Latane and Jones (1977) reported **quarterly standardized unexpected earnings were only reflected in stock prices with a lag**

Givoly and Lakonishok (1979) concluded that **revisions in analysts forecasts of earnings could be used to earn abnormal returns for up to 2 months following release**

Rosenberg, Reid and Lanstein (1985) reported **book to price ratios could help investors to exploit pricing errors after controlling for several ‘risk indexes’**

All of these ‘market anomalies’ could be exploited by investors to **earn a return above that indicated by CAPM**

FACTOR INVESTING

Ross (1976) developed a **theory of security pricing with multiple factors**

The arbitrage pricing theory (APT) was a general framework **which did not specify how many factors would be appropriate**, nor label what the factors might be

In practice, Fama French (1993) 3 factor model relating stock returns to market beta, size and market to book ratio was **first large scale operational model for factor investing**

In another shift from typical investment management practice of the time, implementation of FF 3 factor model is **dollar neutral** (involving both long and short positions) rather than being long only

FLOODGATES OPENED AFTER FAMA FRENCH 1993

After FF (1993), **floodgates for factor investing seemed to open**

Jagadeesh and Titman (93) was first to document **canonical momentum factor** over 3 to 12 month time horizons (buy past winners and sell past losers)

- Geczy and Samonov (2016) documented **momentum in more than 2 centuries of price data**
- Carhart (1997) suggested **momentum should be considered a common factor, leading to 4 factor model**

De Bondt and Thaler (1965) reported opposite of momentum at longer (3 to 5 year) horizons, specifically that **past losers become winners.** They suggest that **stock market overreacts to unexpected news events,** and this can account for a return differential between winners and losers of 25% over 3 years

Momentum and reversion are often **characterized as a result of investor sentiment,** with De Bondt and Thaler (1965) being considered one of first works in behavioural finance

FAST FORWARD TO 2024

As of 2024, there **are over 300[1] market anomalies** documented in the academic literature, across equities, as well as other asset classes such as futures, FX, options and fixed income

Corresponding to this has also been a **large inflow of assets under management** into quantitatively driven strategies

The investment style is currently at a cross roads as it **confronts alpha decay on one hand (due to influx of AUM) and the need to demonstrate superior, or uncorrelated performance** with other common investment styles

[1] https://academic.oup.com/rfs/article/29/1/5/1843824

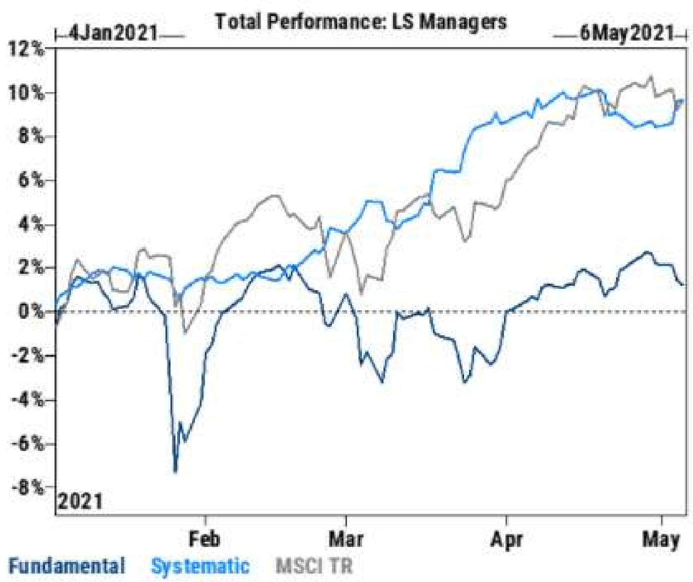
COMMONLY WATCHED FACTORS SORTED BY LAST 1-YEAR RETURN. NOTE MANY FACTORS ARE POSITIVE YTD DESPITE THE OVERALL MARKET BEING DOWN

	Style	Factor/Driver Name (17)	1D Ret	1W Ret	Prior Month
	Highlights	<Filter>			
1)	Growth	5Y Actual Sales Growth	0.56%	1.06%	1.41%
2)	Sentiment	Sell Side Expected Return	0.55%	-1.27%	18.34%
3)	Momentum	PORT US Momentum	0.72%	7.53%	-5.44%
4)	Revisions	3M Target Price Change %	0.37%	10.25%	-6.87%
5)	Growth	1Y Fwd EPS Growth (FY) %	-0.43%	3.42%	-8.00%
6)	Surprises	EPS Surprise % (Last)	-0.16%	0.58%	-0.12%
7)	Size	PORT US Size	-1.85%	-1.38%	1.31%
8)	Revisions	3M EPS Revision % (FY1)	0.11%	6.07%	-6.45%
9)	Volatility	1M Volatility	0.94%	-5.53%	16.51%
10)	Short Interest	SI Days to Cover	0.77%	2.03%	2.51%
11)	Dispersion	Rev Est Dispersion (FY1)	0.54%	-0.43%	9.81%
12)	Share Buybacks	1Y Share Buyback	-0.47%	-3.65%	-1.36%
13)	Profitability	PORT US Profit	-0.77%	-2.72%	-6.02%
14)	Dividends	Dividend Yield (Indicated)	-1.70%	-4.75%	-3.17%
15)	Technicals	14D RSI	0.68%	4.52%	-15.10%
16)	Leverage	PORT US Leverage	-0.05%	-5.68%	8.56%
17)	Value	PORT US Value	-0.85%	-8.23%	-1.19%

Net long-short return of equal weighted stocks in Q1/Q5. Values are negative if bottom

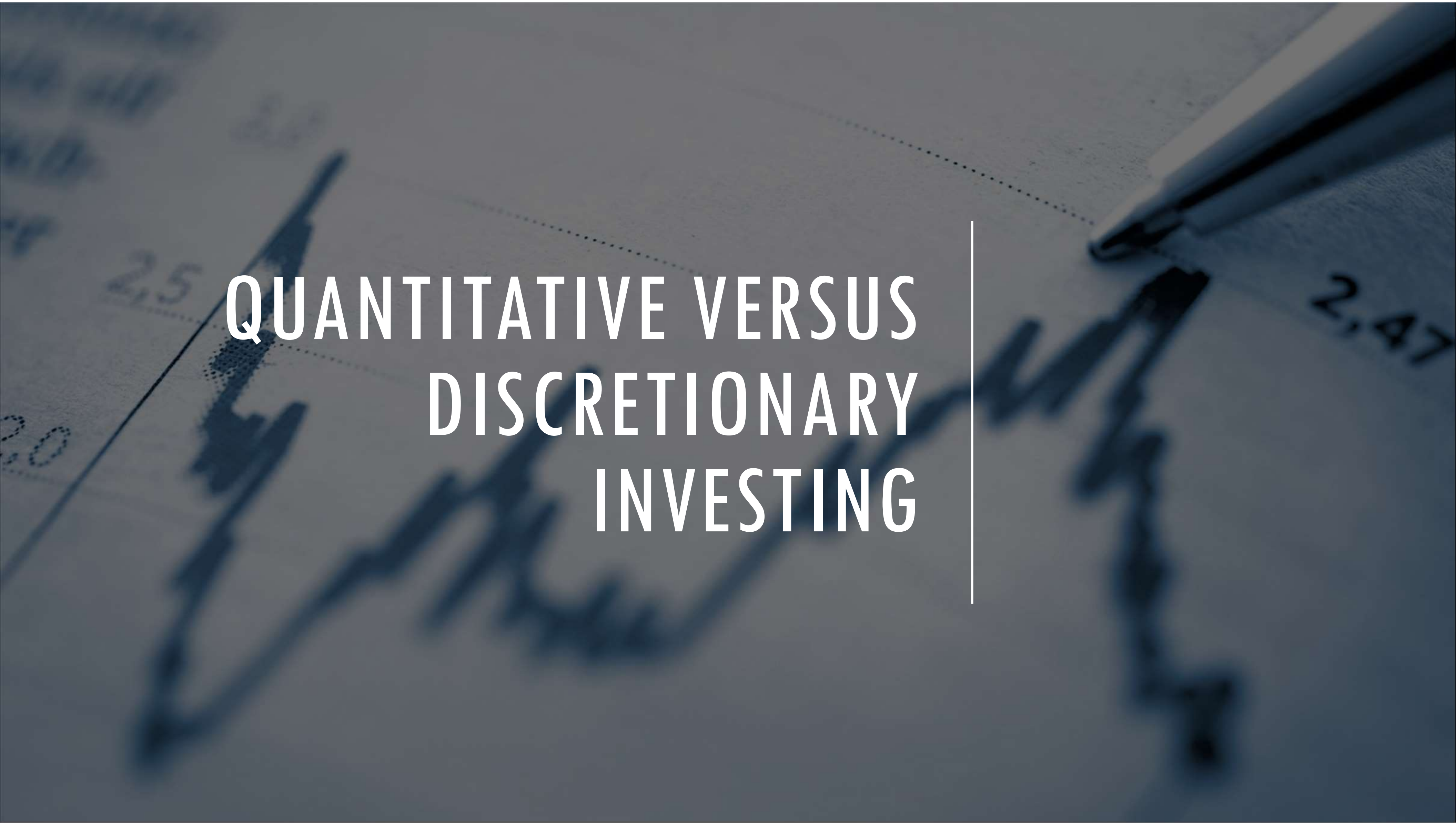
Suggested Functions WB Compare sovereign debt data on one screen VRUS Tra

- **Fundamental LS +1.3% (alpha -6.7%)** vs MSCI TR +9.6%
- **Systematic LS +9.6% (alpha +10.2%)**



1. Most recently, systematic LS has significantly outperformed fundamental LS since 1Q2020
2. Prior to this, fundamental LS had a good run from early 2019 to 1Q2020
3. Both investing styles can coexist together in a single investment portfolio due to low correlations

Source for all graphs and tables: Goldman Sachs Global Markets data as of 6-May-21. Past performance is not indicative of future results.



QUANTITATIVE VERSUS DISCRETIONARY INVESTING

“LAW OF LARGE NUMBERS”

Systematic quantitative strategies are **typically based on statistical methods** to determine whether a specific trade or position is expected to be profitable

Consequently, diversifying portfolio AUM across largest possible tradable universe tends to result in greater statistical validity in backtesting, as well as **more robust live trading performance**

Specifically, by running simulations on the largest possible universe, **power of the simulation is boosted,** leading to higher sharpe ratios.

- This is similar to how **increasing sample size leads to higher t-statistics in estimations**

SYSTEMATIC QUANTITATIVE STRATEGIES TYPICALLY HAVE HIGHER TURNOVER THAN DISCRETIONARY PORTFOLIOS

To achieve statistic validity, we may increase sample size via adding instruments to our portfolio. We may also do so by **decreasing the trading horizon**

Moreover, computer based quantitative trading strategies are typically able to react faster to new information than discretionary strategies. **Trading speed itself may be a source of additional portfolio returns**

Consequently, **quantitative trading strategies generally have higher trading turnover than discretionary portfolios.** As a result, trading costs (e.g. commissions, bid-ask spreads, financing) affect performance more significantly

HIGHER LEVERAGE RATIOS ARE POSSIBLE WITH KNOWN RETURN/DRAWDOWN DISTRIBUTIONS

Many quantitative strategies may individually **have lower % returns than the long term averages** from (say) buy and hold strategies on the broader market

However, they **may have superior risk adjusted return characteristics,** for e.g.

- Lower maximum drawdown relative to buy and hold, and lower volatility
- Very low correlation to buy and hold

In this case, we can **exploit better risk adjusted return characteristics** (e.g. low drawdown) of quantitative trading strategies by levering up so that total expected returns exceed that of market returns, while maximum drawdown does not exceed that of the market

Systematic	Discretionary
Less concentrated positions	More concentrated
Higher trading turnover (sensitive to trading costs)	Lower trading turnover
Generally more highly levered (sensitive to financing costs)	(On average) less highly levered

SYSTEMATIC VERSUS DISCRETIONARY PORTFOLIO MANAGEMENT

SYSTEMATIC STRATEGIES CAN BE CATEGORIZED ALONG TYPE OF INSTRUMENTS AND HOLDING PERIOD

By instrument:

- **Quantitative equity statistical arbitrage** largely construct equity portfolios (often market neutral) based on various market anomalies
- **Asset allocation type strategies** commonly involve portfolio allocation between equities and fixed income instruments
- **Global macro strategies** invest based on economic and political views and their effect on various equity, fixed income, currency, commodities and futures markets

By holding period:

- **>=1 day**: Equities based statistical arbitrage strategies commonly involve daily or monthly rebalance frequencies (depending on the strategy)
- **>=1 quarter**: Global macro / quantitative portfolio allocation strategies may have holding periods that involve quarterly or even longer rebalance
- "**Real time**": Infrastructure sensitive strategies have very short holding periods (sub-second), and frequently involve market making, or arbitrage type trades

IN THIS COURSE, WE WILL SURVEY THE FOLLOWING CATEGORIES OF STRATEGIES

Equities: Medium frequency statistical arbitrage

We will first overview the 'genesis' of statistical arbitrage with pairs trading

Following this, we will survey a selection of market anomalies from the academic literature, as well as replicate a number of these strategies via Python in an open source framework

Many of the econometric tools we survey in this section can be applied directly to other asset classes

Global macro

Quantitative asset allocation via risk parity

Futures curve carry

Trading volatility via VIX futures and options

We will also discuss replication selected strategies

Infrastructure dependent strategies
(high frequency / HFT / ETF arbitrage)

STATISTICAL ARBITRAGE: MARKET NEUTRALITY

Medium Frequency Equities

DOLLAR NEUTRAL, ZERO BETA AND MARKET NEUTRALITY

In equities based statistical arbitrage, we typically construct forecasts or signals based on individual stocks.

Such forecasts / signals may generally be constructed from a wide variety of sources: price-return data, panel data on corporate fundamentals, global economics, natural language processing, behavioural finance, machine learning, etc. Regardless, they are usually specific to a single stock.

How do we combine the various signals into a single portfolio? Objective: we wish to construct a portfolio with zero systematic risk

Why do we want to do this?

- First, investors / clients may not compensate us for taking systematic risk. They can do it themselves
- Second, our strategies may often be silent on making predictions about the overall market
  - Many of our trading algorithms (e.g. pairs trading) make predictions regarding the behaviour of individual securities, or the behaviour of individual securities relative to other individual securities
  - In many cases, it is possible that nothing in our trading strategy makes any prediction about *what the overall market is going to do*
  - In this situation, we may wish for the performance of our portfolio to depend solely (or as much as possible) on the specific spread / variable that we are conditioning on
  - Any PNL / losses due to overall market movements is *accidental* if our strategy does not make explicit predictions in this area

w = [0.25, 0.1, 0.35, 0.3]

Step 1: demean(w)
mean(w): 0.25
demeaned(w) = [0.25 - 0.25, 0.1 - 0.25, 0.35 - 0.25, 0.3 - 0.25] = [0, -0.15, 0.1, 0.15]
norm(demean(w)):
sum(abs(w_1)) = 0.4
final target position vector = [0/0.4, -0.15/0.4, 0.1/0.4, 0.15/0.4] = [0, -0.375, 0.25, 0.375]

DOLLAR NEUTRAL, ZERO BETA AND MARKET NEUTRALITY

Consider the specific economic environment in (say) **2Q 2020**. Many economies are ‘reopening’ after coronavirus induced shutdowns

Taking perspective of a portfolio manager on **1 May 2020**, we have following (hypothetical) binary outcomes:

1. Global economy reopens over next several months and there is no second infection wave. Demand and production recover simultaneously. Coupled with unprecedented fiscal and monetary support measures, this scenario is consistent with equity markets reaching new highs

2. Global economy reopens over next several months, and there is consistently second infection wave in countries that reopen. Closing is mandated again to save lives. Equity markets conclude that sustained opening is predicated on discovery of a vaccine. Due to uncertainty and time to operationalize a vaccine, markets hit multi-year lows

Although differentiating both scenarios above are first order important, many trading strategies currently existing make no prediction about their likelihood. Rather, many of the strategies we implement will forecast for a single stock only. Hence, by taking ‘zero market risk’, we insulate the performance of our strategy from difficult to forecast scenarios

To be blunt, back in 2020, it may not have been possible to win a quantitative trading strategy (using classical or machine learning methods) to forecast between both scenarios

DOLLAR NEUTRAL, ZERO BETA AND MARKET NEUTRALITY

Example two: Fast forward to **2Q 2022**. There is a war in Ukraine, rising inflation and a potential left tail event in French presidential elections [*extreme* left tail: France may leave Euro if Le Pen wins with a large enough margin]

Taking perspective of a portfolio manager on **1 April 2022**, we have following possible outcomes:

1. Peace in Ukraine, results in falling commodity prices. Coupled with post Covid-19 reopening, this leads to untangled supply chains and hence falling inflation. Monetary tightening becomes less of a priority. Macron wins the French presidential elections with a good margin, and the Euro is safe after all

2. Nuclear or chemical escalation in Ukraine, new Covid-19 strain results in reimposition of lock-down measures, France's position in the Euro looks shaky ….

Similar to before, many of our strategies forecast individual instrument performance, not the outcome of these events

It may be impossible (given the current state of data availability and technology) to forecast any of the above.

Again, by taking 'zero market risk', we insulate the performance of our strategy from difficult to forecast scenarios

DOLLAR NEUTRAL, ZERO BETA AND MARKET NEUTRALITY

A zero beta portfolio can be constructed over a set of stocks S1, S2, …Sn, by choosing weights such that:

$\sum_i w_i \beta_i = 0 \; where \; w_i \; are \; portfolio \; weights$

In practice, we commonly approximate this by creating a dollar neutral portfolio instead, where sum of all notional values = 0. i.e.

$\sum_i w_i = 0 \; where \; w_i \; are \; portfolio \; weights$

INTRODUCTION: STATISTICAL ARBITRAGE – DOLLAR NEUTRAL

For the rest of this class, we will use the phrase ‘dollar neutral’ to refer to zero net portfolios

Dollar neutral framework has been extended from pairs trading to portfolio formation based on various corporate fundamentals, accounting, sentiment and other variables

There is a rich strand of the academic literature that focuses on market anomalies (e.g. Fama French factors, momentum, earnings quality, analyst revision momentum).

Most of these anomalies are implemented in a market neutral framework, both in the academic literature as well as industry

CONSTRUCTING A DOLLAR NEUTRAL PORTFOLIO

Note that we can transform the output of any equities trading strategy to form a market neutral portfolio

A portfolio is defined as a set of weights that allocate investment capital. A trading strategy is defined as an algorithm that generates a set of portfolio weights at regular time intervals

E.g. Portfolio weights W = [w_1,…,w_n] for a portfolio with n assets

Without loss of generality, assume $\sum_i abs(w_i) \quad = 1 \; where \; w_i \; are \; portfolio \; weights$

This implies that each of w_1 …w_n are % allocation of the portfolio's total capital

- E.g. w1 = 0.01 means that there is a long position on asset 1 with notional value 1% of portfolio's capital
- W3 = -0.02 means that there is a short position on asset 3 with notional value 2% of portfolio's capital

DOLLAR NEUTRALITY CAN BE APPLIED TO ANY EQUITIES TRADING STRATEGY

Assume we have equities trading strategy f(x) that produces for each trading period a set of weights $[w_1, \dots w_n]$

There are no restrictions on the weights except that they be real numbers. They can be positive or negative

We can transform this portfolio into a market neutral portfolio by demeaning this vector

i.e. transform weight vector into weights $[w'_1,\dots,w'_n]$ by:

$w'_i = w_i - w_{\_mean}$ where $w_{\_mean}$ is simple average of all weights

for each grouping in the data:
step 0: compute portfolio vector just for positions in that group: w_sectorA = [0.1, 0.2, 0.3]
step 1: mean of vector from step 0
step 2: normalize vector from step 1

INDUSTRY NEUTRALITY

An extension to market neutrality is industry neutrality.

We can convert the output of any quantitative strategy (portfolio weights $[w_1, \ldots w_n]$) into an industry neutral format by subtracting the industry mean from each individual weight (instead of the 'global' mean, as in previous slide)

i.e. transform weight vector into weights $[w'_1,\ldots,w'_n]$ by:

$w'_i = w_i - w_{\_industry\ mean}$ where $w_{\_industry\ mean}$ is industry mean of the weights

[PRACTICUM INTERLUDE]

For this section, we will transition to demonstrating the effect of market and industry neutrality on a simple mean reversion strategy on 2000 companies simultaneously (Russell 2000)

We will be running this on panel data and Python code which has been uploaded to eLearn folder

Uploaded Python codebase also contains an embedded back-testing simulator which has been built from first principles using only numpy and pandas libraries. You may be able to adapt this in future for your own purposes [e.g. final project, etc]

STATISTICAL ARBITRAGE: PAIRS TRADING

Medium Frequency Equities

INTRODUCTION TO PAIRS TRADING

Statistical arbitrage on equities can **trace its origins to market neutral pairs trading pioneered by Morgan Stanley in the 1980s**

Pairs trading in its original form **involves identifying two companies (say Pepsi and Coca Cola) which should fundamentally have very similar businesses.**

Whenever there are short term deviations in returns between the two instruments, we can **go long the instrument which is temporarily undervalued, and short the instrument which is temporarily overvalued**

The assumption is that temporary "spread" between both instruments should eventually converge, and because we entered the pair at a "positive spread", **we should realize a profit as the spread converges back to 0**

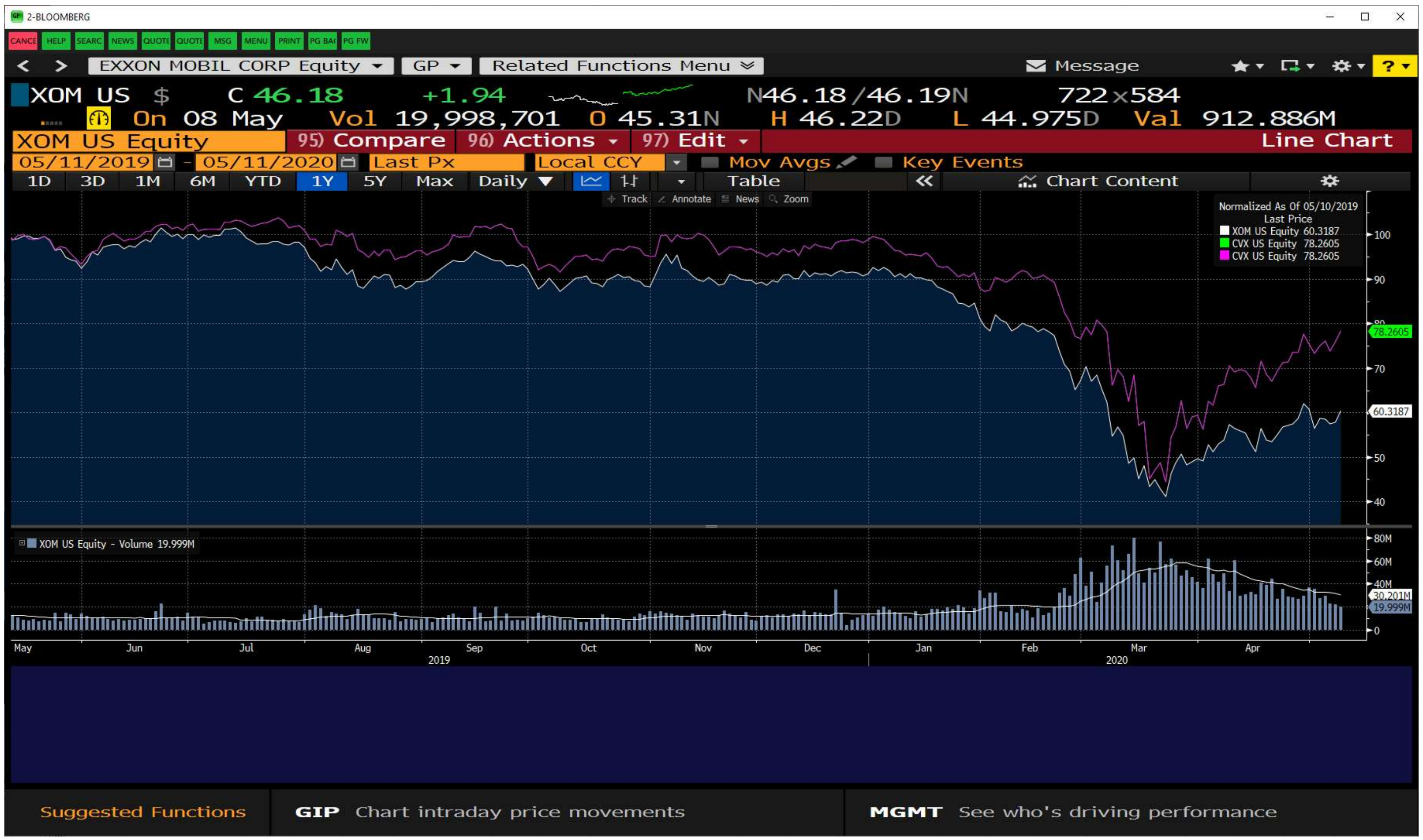
PEPSI VERSUS COCA COLA (2019 TO 2020)

Relationship between Pepsi and Coca Cola has held up pretty well until 1Q 2020.

If we think the spread will mean revert, we can short the stock that is trade above (currently Pepsi) and go long the stock that is trade below (Coca Cola)

A sharp market downturn or upturn should impact both legs of the trade. Hence, this strategy may profit in both states

EXXON VERSUS CHEVRON (2019 TO 2020)



Note that the relationship between both stocks do not need to be 1:1

For e.g. in this chart, we observe that Exxon consistently trades slightly above Chevron

The optimal ratio may then be 0.9 Exxon: 1 Chevron (for e.g.)

PAIRS TRADING 101

In this example. assume **Pepsi and Coca Cola are trading at USD$100.** In today's trading session, Pepsi increases to USD$102, while Coca Cola falls to USD$98. The spread between both stocks is USD$4

In this situation, we would short 1 share of Pepsi at USD$102, and buy 1 share of Coca Cola at USD$98. At the inception of this trade, we generate positive cash of USD$4.

- Our short position on Pepsi generates positive cash of USD$102. By 'shorting', we are in effect "selling" Pepsi stock, and we get USD$102 for each share of Pepsi
- Our long position on Coca Cola requires us to pay USD$98, since we are 'buying' the stock.
- **Total positive cash generated at trade inception is USD$4.**
- We also have equity positions on Pepsi and Coke which are [-1, 1] shares. This is our portfolio at the end of today's trading session.

PAIRS TRADING 101

In tomorrow's trading session, Pepsi falls back to USD$101, while Coca Cola rises to USD$101. The 'spread' between both stocks is now USD$0. **We now close both legs of the trade, which (on balance) does not require any additional cash outflows.** Thereafter, we get to keep the USD$4 we made initially, while having no additional equity positions

- Closing our short position on Pepsi requires "buying' 1 share of Pepsi at USD$101. This requires cash outlay of USD$101
- Closing our long position on Coca Cola requires "selling" 1 share of Coca Cola at USD$101. This generates cash receipts of USD$101
- **There is therefore no *net* cash outlay or inputs in tomorrow's trading session**
- Neglecting trading commissions, short borrow costs and financing costs, we have therefore made a total profit of USD$4 over 2 trading days.
- What is percentage profit?

PAIRS TRADING 101

In the previous example, our **equities portfolio involved shorting 1 share of Pepsi at $102, and buying 1 share of Coke at $98**

Total *Gross Market Value (GMV)* of equities portfolio **is $200 at inception**. Gross market value is absolute sum of all long and short positions

Total *Net Market Value (NMV)* of equities portfolio is **$-4 at inception.** Net market value is sum of all long and short positions

- For our short position, we 'owe' the broker 1 Pepsi share, which was trading at $102 at inception. Hence, this reduces our equities portfolio value by -$102 [*more on next slide*]

For portfolios with short positions, returns are typically computed over GMV. In this case, our % returns **is $4/200 or 2%**

PAIRS TRADING 101: MECHANICS OF SHORT SELLING

In practice, a short trade has 2 components:

- A **conventional 'sell trade' which generates cash in the portfolio.**
  - The prime broker may pay interest on cash in the portfolio. This may, for e.g. be US 3 month LIBOR – spread. Both the value of the spread as well as benchmark is an outcome of bargaining with the prime broker

- **A stock loan trade which provides the equity to sell**; stock loan invites a short borrow fee
  - "General collateral stocks" in the US have short borrow fees around 35 bps / year

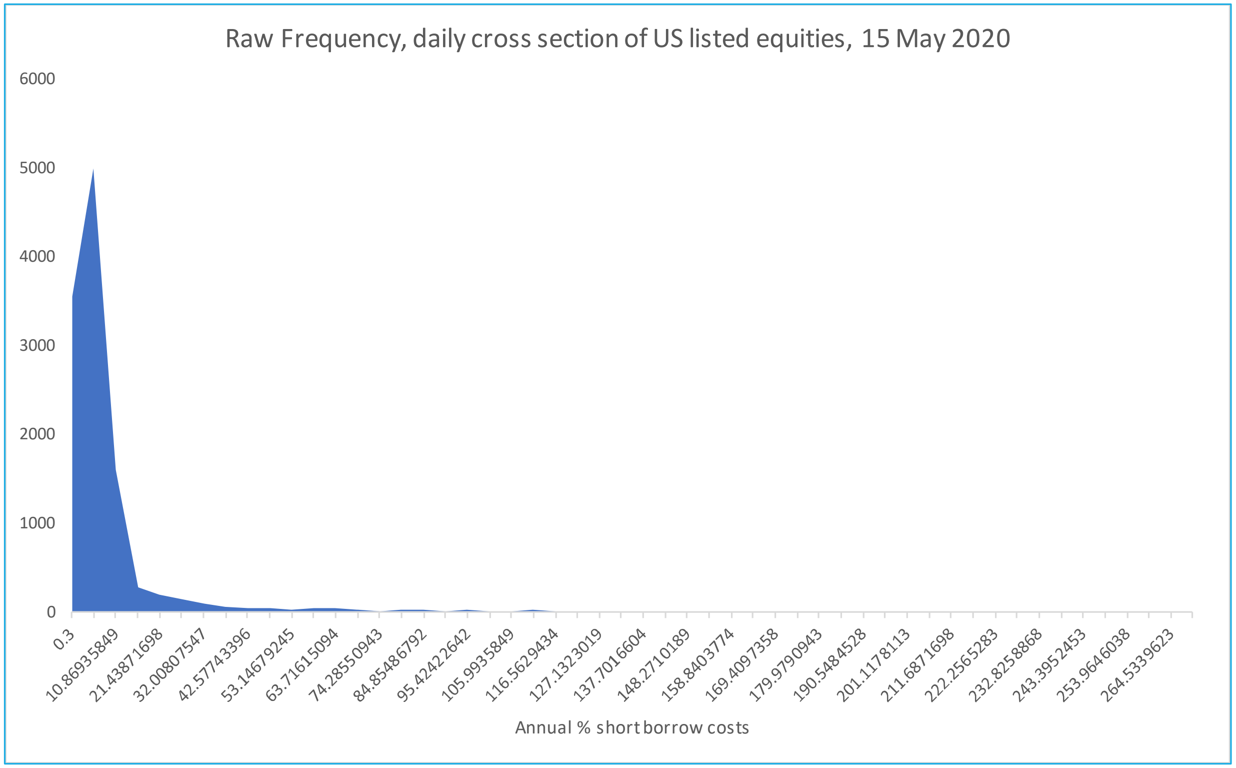
For instance, on Pepsi short example, **we initially started with 0 pepsi shares in the portfolio.**

1. Our prime broker helps us source for 1 pepsi share to short, and lends this share to us. After this interim and transient step, we now have 1 pepsi share in our portfolio, and we 'owe' the prime broker 1 pepsi share
2. We now sell that 1 pepsi for USD$102. After this final step, we now have USD$102 cash in the portfolio (from selling), and we still owe our prime broker 1 pepsi share. The value of that 'loan' is the market price of Pepsi, which would be USD$102 in this cash.
3. Note that if there are no changes in the share price of Pepsi, there is no net change in the value of our portfolio. We got USD$102 in cash, and we owe USD$102 to the prime broker.
4. If we were to hold our short position for an extended period of time, the short borrow fee would then accrue on the market value of our loan with the prime broker, whilst cash interest would be credited on the cash generated in our portfolio
5. In practice, steps (1) and (2) occur almost simultaneously during a short trade

PAIR TRADING 101: MECHANICS OF SHORT SELLING

Short locates:

- In the case of Pepsi-Coke, being able to 'locate' sufficient shares to short is generally not a concern
- There may be case of 'hard to locate' stocks with limited inventory for shorting
- In practice, short borrow fees can exceed 100% p.a. for hard to locate stocks
- Most brokers make it a requirement to pre-locate stocks for shorting before entering a short trade
- In practice, we should avoid constructing pairs where either leg is based on an illiquid or hard to short stock
- Any backtest based on such instruments are likely 'too good to be true'



SUMMARY STATISTICS ON ANNUALIZED SHORT BORROW COSTS FOR US LISTED EQUITIES

PAIRS TRADING REVISITED

Returning to our Pepsi-Coke example, here are some questions before we can systematize a Pairs Trading strategy

- What is statistical / econometric criteria to identify all pairs over a liquid equities universe (say S&P 500 or S&P 100)?
- Do the relationship between both legs in a pair need to be 1:1? If not, how can we calibrate exchange ratio?
- How do we know when to exit? Should we have a stop loss?
- How do we know how much to invest in each pair (if we have multiple pairs)?

For the rest of today's session, we will focus on answering these questions

PAIR IDENTIFICATION

There are multiple widely used methodologies for pair identification:

1. Correlation
2. Cointegration
3. Distance measure

At a high level, we may iterate over all possible pairs in our trading universe, and evaluate each pair with any / all of the above. Each methodology will provide diagnostic statistics that evaluate 'goodness of fit'

One takeaway from this list is that methodologies for pair identification is by no means a "closed book", and new methodologies are continually added to the list by researchers and portfolio managers

PAIR TRADING WITH CORRELATIONS

Medium Frequency Equities

CORRELATION AS A DISTANCE MEASURE

We may also compute the correlation between the price series of both legs of a candidate pair to determine if they are a good fit

Formally, given price series X1t and X2t, we compute over a rolling window (e.g. 2 years):

$\text{Correlation}(X_{1,t}, X_{2,t}) = \text{COV}(X_{1,t}, X_{2,t}) / \text{SD}(X_{1,t}).\text{SD}(X_{2,t})$

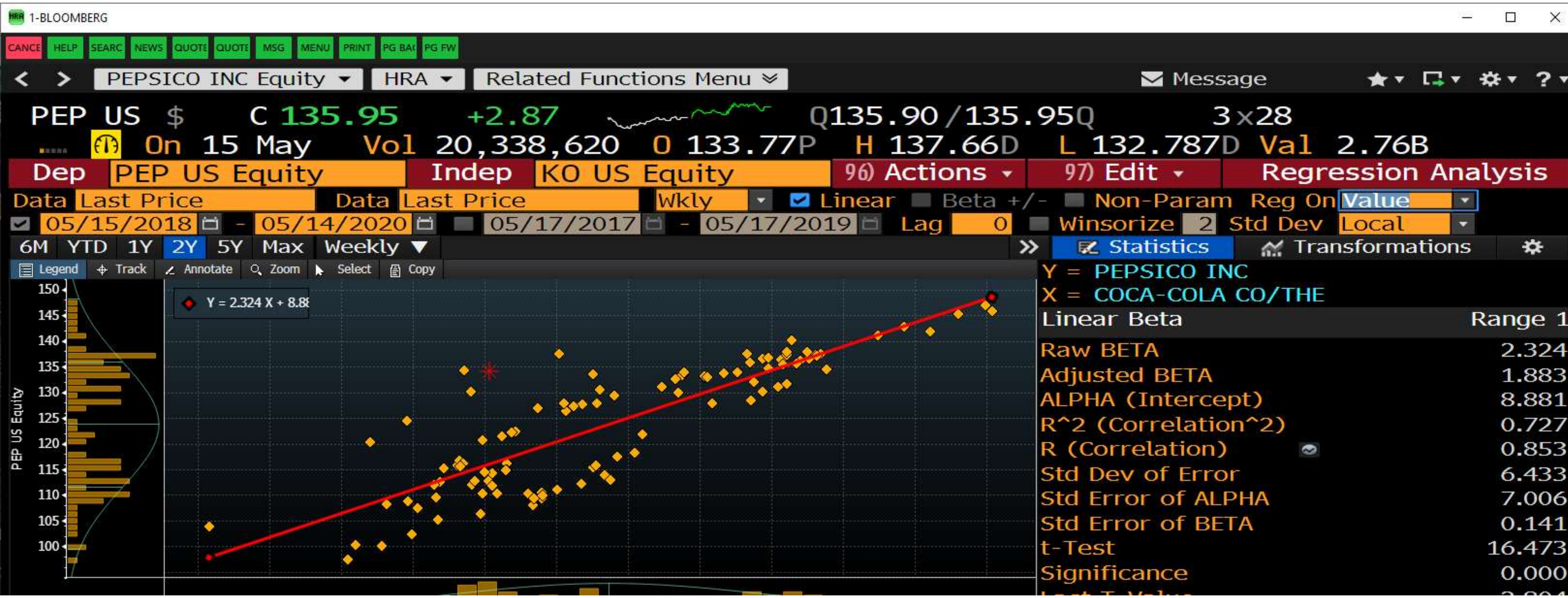
We may then shortlist pairs where both legs exhibit a correlation above a cutoff (e.g. 0.9).

CORRELATION AS A DISTANCE MEASURE: SOLVING FOR HEDGE RATIO

We note that an OLS regression of $X_{1,t}$ against $X_{2,t}$ would simultaneously allow us to estimate the hedge ratio (i.e. how many shares of $X_{1,t}$ to trade against 1 share of $X_{2,t}$), as well as goodness of fit of the pair

This is because R^2 of an OLS regression with a single independent variable and intercept term is also equal to the squared Person correlation coefficient of $X_{1,t}$ and $X_{2,t}$

Hence, we can run OLS regressions of X1t against X2t over all possible pairs in the universe, and shortlist pairs with high R_square. The OLS beta will be the hedge ratio



OLS REGRESSION OF PEPSI'S STOCK PRICE VERSUS COKE

2Q 2019 – 2Q 2020

CRITIQUE OF OLS METHODOLOGY

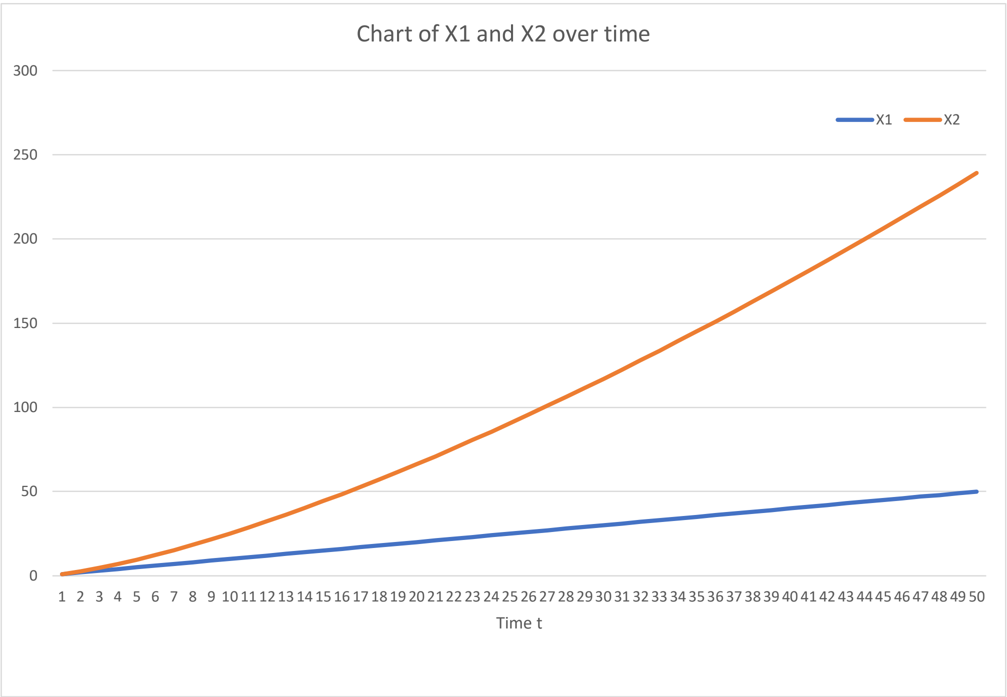
OLS does not make any statement on whether spread ($X_{1,t}$-$\beta X_{2,t}$) is stationary or not

In case of non-stationary spread, it is still possible for OLS to give an extremely high R^2, although (to be blunt), you can lose lots of money trading on this

HERE IS NEGATIVE EXAMPLE ...

NEGATIVE EXAMPLE: NON STATIONARY SPREAD WITH HIGH R^2

Consider this **negative** example: $X_{2,t} = X_{1,t}^{1.4}$



What will be R^2 of OLS regression of $X_{2,t}$ against $X_{1,t}$?

NEGATIVE EXAMPLE: NON STATIONARY SPREAD WITH HIGH R^2

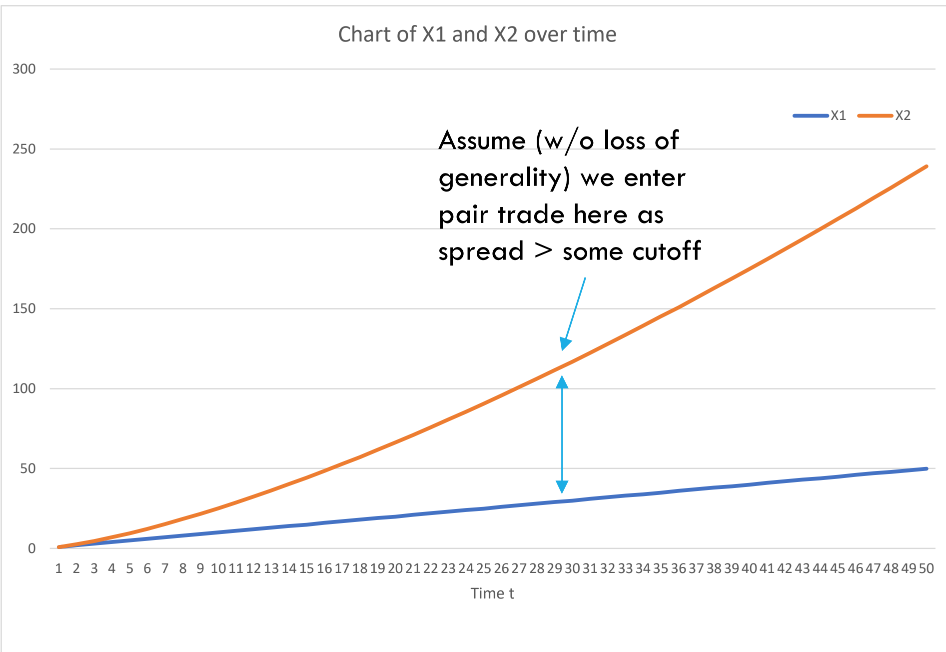
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.993452309							
R Square	0.98694749							
Adjusted R Square	0.986675563							
Standard Error	8.417708468							
Observations	50							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	257175.1646	257175.1646	3629.453738	6.8937E-47			
Residual	48	3401.175161	70.85781585					
Total	49	260576.3398						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-24.70792843	2.417054946	-10.22232799	1.22693E-13	-29.56774311	-19.84811374	-29.56774311	-19.84811374
X Variable 1	4.969778192	0.082492862	60.24494782	6.8937E-47	4.803915176	5.135641209	4.803915176	5.135641209

Very high R^2.
> 0.98.
Superficially this looks like good candidate for pair trading

This tells us that we should short around (1 / 4.97) shares of X2 for each share of X1 we buy

NEGATIVE EXAMPLE: NON STATIONARY SPREAD WITH HIGH R^2

Despite high R^2 and apparently reasonable hedge ratio, this pair will never converge.



Note that after trade entry (where we short X2, orange line and long X1, blue line), we will consistently lose money faster on the short leg than we make on the long leg. The spread keeps increasing without limit

CRITIQUE OF OLS METHODOLOGY

Constructively:

1. We can use OLS methodology as a 'consistency check' in concert with other pair identification methodologies. E.g. if multiple methodologies agree on a pair, it might be more robust

2. Implementing a 'stop loss' routine may help us to leg out of pairs that have non-stationary spreads, hence limiting downside on those pairs

3. We may greatly improve upon the OLS approach by checking for stationarity of the spread …

PAIR TRADING VIA COINTEGRATION

Medium Frequency Equities

REVIEW OF COINTEGRATION

- Definition: Cointegration is a property of 2 or more time series which means that a linear combination of the variables is stationary
- A stationary time series is one whose statistical properties such as mean, variance, autocorrelation, etc are constant over time
- This is a desirable property. We would like the 'spread' between the time series of S1 price and S2 price to remain constant over time.
- Any deviation is (by definition) temporary and transient, and will therefore revert eventually to long run mean
- If two time series are cointegrated, we can model them with an error correction model, or ECM
- If more than two time series are cointegrated, we can use a vector error correction model, or VECM

Background on Cointegration and Error correction

Introduction to ECMs and VECMs

- Error Correction Models allow us to **model nonstationary data directly without differencing**
  - Usually both variables will be I(1), which means they are stationary after 1 difference.
  - Only requirement to using ECM is **that x variable be cointegrated with y variable**
- There are two main forms of ECMs:
  - **Engle Granger ECM**. This version admits only a single x variable to forecast the y variable
  - **Johansen VECM**. This implementation is able to model multiple y variables with multiple x variables
- Code implementation for this class **will focus on Johansen VECM**. This is because the VECM is a superset of ECM, and you can always express a 2 variable system using the multiple variable syntax easily

ECM definition (1 variable)

Consider following structure where we wish to forecast Y_t with x_t

- $\Delta Y_t = \underbrace{\gamma + B_1 \Delta X_t}_{\textit{Equilibrium}} + \underbrace{\alpha(Y_{t-1} - B_0 - B_1 X_{t-1})}_{\textit{Transient}} + v_t$

- **Two segments** to equation above.

- **Equilibrium** (long run) term relates the change in Y with change in X. E.g. a 1 unit change in X results in B_1 units permanent change in Y

- **Transient** term relates the change in y with deviation from long run relationship between Y and X from previous period

Error Correction Model (ECM) definition with 1 variable

Consider following structure where we wish to forecast Y_t with x_t

- $\Delta Y_t = \gamma + \underbrace{B_1 \Delta X_t}_{\textit{Equilibrium}} + \underbrace{\alpha(Y_{t-1} - B_0 - B_1 X_{t-1})}_{\textit{Transient}} + v_t$

- **Focusing on transient relationship part of the ECM:**
  - We hypothesize long run relationship is given by OLS equation: $\mathbf{Y_t = B_0 + B_1 X_t + e_t}$
  - $\mathbf{Y_t - B_0 - B_1 X_t = e_t}$ **(white noise)**
  - $\mathbf{E(Y_t - B_0 - B_1 X_t) = 0}$
  - In practice, gap $(Y_t - B_0 - B_1 X_t)$ **may temporarily exhibit large values**
  - If gap temporarily widens, it **should narrow over time and revert back to 0**
  - For example, if Y_{t-1} is abnormally high in period t-1, such that $\mathbf{Y_{t-1} - B_0 - B_1 X_{t-1} = C >> 0}$, we will predict Δ**Y will be lower by** α**C in current time period** ($\alpha < 0$), to bring Y back in line with "long term" trends

Engle Granger ECM estimation with single X variable in R

- Given a pair of variables, X_t and Y_t, determine if **they have the same order of integration**
- If so, **estimate "trial" cointegrating relationship between the two variables using OLS**
- **Test residuals from "trial" OLS estimation** for a cointegrating relationship
- Depending on output from above steps:
  - If a cointegrating relationship exists between both variables, **estimate an ECM as the "best" predictive model for Y in terms of X**
  - If no cointegrating relationship exists and **both variables are I(0), use simple OLS**
  - If no cointegrating relationship exists **and both variables are I(1) or above**, use OLS on the differenced variables
- (Note: we are currently not discussing having lagged values of Y on the RHS. **This will be discussed under Vector ECMs**)

Recall Test for Cointegration

- We can test two variables for cointegration using the **Engle Granger test**

- **Engle Granger test** performs the following:

    - **Performs static OLS regression**, and collects residuals from this regression

    - Test residuals for existence of unit root using **Augmented Dickey Fuller (ADF) test**

    - If both time series are cointegrated, then the **residuals will be judged to be stationary as a result of the ADF test**

VECM with multiple variables

- So far, we have considered cointegration between one pair of variables only.
  - With just two variables in the system, it is only possible to have only a maximum of 1 cointegrating relationship.
  - Hence, if just studying a system with 2 variables, Engle Granger ECM is sufficient

- What about more general cases where we are studying relationship between N variables?

- Given an arbitrary group of N variables, there are a maximum possible of N-1 linearly independent cointegrating relationships

VECM with multiple variables

- We will study multiple cointegrating relationships in the context of a VAR model of order p, with N variables being modelled
- For example, we can build a VAR with 3 variables to model consumption (C_t), income (I_t) and savings (S_t). Order (number of lags) for the VAR will depend on the output from VARselect()
- In this setting, there might be up to two possible *linear independent* cointegrating relationships, e.g
  - $C_t = B_0 + B_1 I_t + B_2 S_t + e_{1,t}$
  - $I_t = B_3 + e_{2,t}$

VECM with multiple variables

- In the VECM methodology, we insert error correction terms to the VAR specification. Primary steps in the methodology are:
- Specify and estimate a VAR(p) model for N variables
- Determine the number of cointegrating vectors via maximum likelihood
- If there are 0 cointegrating vectors, we cannot build a VECM model. Proceed to render all variables stationary, and build a VAR on the differenced variables (as per before)
- If there is at least 1 cointegrating vector:
  - We can build a VECM
  - Also, we should build a VECM. In the presence of cointegration, model will be a better description of data compared to VAR on differenced variables
  - We can estimate VECM by maximum likelihood

Example: VECM as a model for PPP

- Consider purchasing power parity (PPP) between 2 countries under floating exchange rates
- Purchasing power parity states that the same basket of goods should cost the same in different countries
- i.e. price of basket in country A (P^1) = price of same basket in country B (P^2) exchange rate (X)
- In practice, economists rely on price indexes whose market baskets differ across countries, so the PPP equation needs a constant of proportionality to reflect this difference.
  - $X = A_0 P^1 / P^2$
  - Ln(X) = Ln(A0) + Ln(P1) - Ln(P2)
- Using small letters to represent logs of variables, we could estimate a VECM (with 1 lag), and 1 cointegrating relationship

Example: VECM as a model for PPP

$$\Delta x_t = \beta_{x0} + \beta_{xx1}\Delta x_{t-1} + \beta_{x11}\Delta p^1_{t-1} + \beta_{x21}\Delta p^2_{t-1} + \lambda_x\left(x_{t-1} - \alpha_0 - p^1_{t-1} + p^2_{t-1}\right) + v^x_t$$

$$\Delta p^1_t = \beta_{10} + \beta_{1x1}\Delta x_{t-1} + \beta_{111}\Delta p^1_{t-1} + \beta_{121}\Delta p^2_{t-1} + \lambda_1\left(x_{t-1} - \alpha_0 - p^1_{t-1} + p^2_{t-1}\right) + v^1_t$$

$$\Delta p^2_t = \beta_{20} + \beta_{2x1}\Delta x_{t-1} + \beta_{211}\Delta p^1_{t-1} + \beta_{221}\Delta p^2_{t-1} + \lambda_2\left(x_{t-1} - \alpha_0 - p^2_{t-1} + p^2_{t-1}\right) + v^2_t.$$

Source: Sims, Christopher A. 1980. Macroeconomics and Reality. Econometrica 48 (1): 1- 48

- If exchange rate is out of equilibrium, we will expect some adjustment back towards the long run equilibrium in the next period
- Error coefficients λ_x, λ_1 and λ_2 measure these responses.
- Using the above logic, we expect λ_x and λ_2 to be negative and λ_1 to be positive

Example: VECM as a model for PPP

- Within PPP framework, **there may also be a second cointegrating relationship**
- Suppose country 1 is on gold standard**. Price level in this country would be constant in the long run**
- i.e. $\mathbf{p_1} = \alpha_1$. This would be second cointegrating relationship

Example: VECM as a model for PPP

- VECM incorporating both cointegrating relationships:

$$\Delta x_t = \beta_{x0} + \beta_{xx1}\Delta x_{t-1} + \beta_{x11}\Delta p^1_{t-1} + \beta_{x21}\Delta p^2_{t-1} + \lambda_x\left(x_{t-1} - \alpha_0 - p^1_{t-1} + p^2_{t-1}\right) + \mu_x\left(p^1_{t-1} - \alpha_1\right) + v^x_t$$

$$\Delta p^1_t = \beta_{10} + \beta_{1x1}\Delta x_{t-1} + \beta_{111}\Delta p^1_{t-1} + \beta_{121}\Delta p^2_{t-1} + \lambda_1\left(x_{t-1} - \alpha_0 - p^1_{t-1} + p^2_{t-1}\right) + \mu_1\left(p^1_{t-1} - \alpha_1\right) + v^1_t$$

$$\Delta p^2_t = \beta_{20} + \beta_{2x1}\Delta x_{t-1} + \beta_{211}\Delta p^1_{t-1} + \beta_{221}\Delta p^2_{t-1} + \lambda_2\left(x_{t-1} - \alpha_0 - p^2_{t-1} + p^2_{t-1}\right) + \mu_2\left(p^1_{t-1} - \alpha_1\right) + v^2_t$$

Source: Sims, Christopher A. 1980. Macroeconomics and Reality. Econometrica 48 (1): 1- 48

- This would be VECM with 3 variables and 2 cointegrating relationships, of order 1

- If we wanted to increase order of VECM, we can simply include more lagged differences in equilibrium portion of VECM

Cointegration applied to Pairs Trading

COINTEGRATION AND PAIRS TRADING

1. In pairs trading, we focus on the transient term in the ECM or VECM, and **assume that error correction will occur**

2. i.e. we assume that at the start of our evaluation period, both time series are in equilibrium (this is an assumption for backtesting to check). Thereafter, we assume that **any deviations from equilibrium will be mean reverted.**

3. This is equivalent to measuring deviations from long run relationship between both stocks (or within a group of stock), and **assuming that they will revert back to the long run relationship**

COINTEGRATION AND PAIRS TRADING

$$spread_t = X_{2,t} - \beta X_{1,t}$$

Assume that we buy one share of stock 2 and sell short β share of stock 1 at time $t-1$. $X_{1,t}$ and $X_{2,t}$ represent the price series of stocks 1 and 2 respectively. The profit of this trade at time t, is given by:

$$(X_{2,t} - X_{2,t-1}) - \beta(X_{1,t} - X_{1,t-1})$$

This would be (after rearranging):

$$(X_{2,t} - \beta X_{1,t}) - (X_{2,t-1} - \beta X_{1,t-1}) = spread_t - spread_{t-1}$$

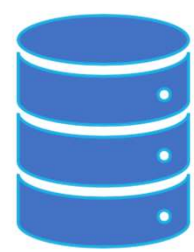
COINTEGRATION AND PAIRS TRADING

The profit of buying 1 share of stock 2 and selling β shares of stock 1 for the period t is given by change in the spread for that period

We know that the spread is stationary by definition and therefore mean reverting

PORTFOLIO FORMATION

We can also start with all possible pairs in the universe (ignoring SSD criteria) and search for cointegrated pairs



TESTING FOR COINTEGRATION IN PORTFOLIO FORMATION

We modify two step Engle-Granger approach to test for the existence of cointegration between candidate pairs, and also to estimate the cointegration coefficient beta (which is also the hedge ratio).

In this procedure, cointegration regression is estimated using OLS in the first step, and then we test for the cointegration relationship using the hedge ratio estimate.

2 STEP ENGLE GRANGER PROCESS

Specifically,

(Zero step) Test that $X_{1,t}$ and $X_{2,t}$ are integrated to the same order I(1)

First run this regression to estimate beta:

regress $X_{2,t}$ versus $X_{1,t}$

Second, given the estimate of beta from step 1, we compute predicted residuals from same regression in Step 1, and used in a regression of differenced variables plus a lagged error term.

We then test for cointegration using a standard t-statistic

TRADING RULES

We enter the pair (simultaneous long and short positions) when the normalized spread diverges beyond 2, where the normalized spread is:

$$spread_{normalized} = \frac{spread - \mu_e}{\sigma_e}$$

Note that we compute Ue by forming the equation below and taking average across all t. Same for sigma e

$$spread_t = X_{2,t} - \beta X_{1,t}$$

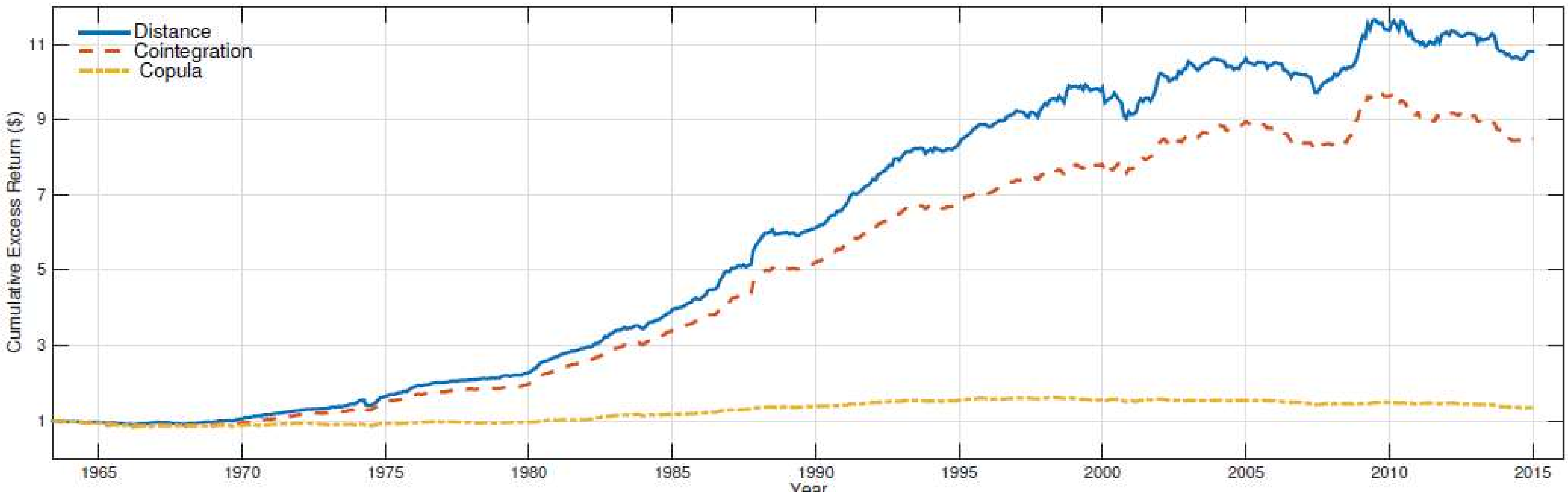
TRADING RULES

If the spread drops below -2, we buy 1 dollar worth of stock 2 and short β dollars of stock 1.

We short $1/\beta$ worth of stock 2 and buy 1 dollar of stock 1 when spread moves above +2

We close both legs once spread returns to 0, which translates into pair returning to its long term equilibrium

This figure shows the evolution of wealth based upon an investment of $1 in each strategy. The return on employed capital after transaction costs is applied for the calculation of cumulative excess returns.



HISTORICAL PNL

We note that corpula approach has historically been another approach to pairs trading

PAIR TRADING WITH A DISTANCE MEASURE

Medium Frequency Equities

PAIR IDENTIFICATION: DISTANCE MEASURE

In this approach, a distance measure between each leg of a pair is formed

Pairs with distance measure below a threshold are grouped into pairs.

We would trade the pair when their distance in live-quoted prices increase, and close the position as the distance falls

PAIR IDENTIFICATION: DISTANCE MEASURE

Starting with the universe of S&P 500 stocks, we may first restrict pair formation to same-industry stocks.

- This is because pairs in the same industry are likely to have more similar economics

Compute distance between two stocks as (Gatev et al):

- Sum squared difference in normalized price between both legs during formation period (12 months)

PAIR IDENTIFICATION: DISTANCE MEASURE ALGORITHM (GATEV '06)

Trading Algorithm:

- At the start of each 12 month period, normalize all stock prices to \$1 by dividing previous period normalized price by the cumulative return in last 12 months (including dividends and corporate actions)
- Using normalized stock prices, compute spread for each possible pair in trading universe as price(stock 1)/price(stock 2) - 1 over the last 12 months
- Compute sum of square deviations between two normalized price series over formation period (SSD)
- For each security in our universe, matching partner is stock that minimizes sum of squared deviations between both normalized price series
- On portfolio formation date, we sort pairs in ascending order of distance (SSD), and chose top-N pairs to trade.
- For each of top-N pairs, we enter a position if the spread in actual trading is more than 2 historical standard deviations from the historical average
- We unwind positions at the next crossing of the prices

PAIR IDENTIFICATION: DISTANCE MEASURE ALGORITHM (GATEV '06) WORKED EXAMPLE

Say we start with the S&P 100. Screen out all stocks that have 1 or more days with no trade (this is highly unlikely with S&P 100)

There are $O(N^2)$ possible pairs of stocks where N is universe size

Consider any two stocks, S1 and S2 in S&P100

PAIR IDENTIFICATION: DISTANCE MEASURE ALGORITHM (GATEV '06) WORKED EXAMPLE

Assume **it is 1 Jan 2020 now** (when we are forming portfolio of pairs to trade).

Assume following evolution of S1 and S2 stock prices from 1 Jan 2019 to 1 Jan 2020. For brevity, we only list a few dates, and we do not factor in dividends (yet)

Date	1 Jan 2019	1 March 2019	1 May 2019	1 July 2019	1 Sep 2019	1 Nov 2019	1 Jan 2020
S1	182	179	185	196	217	198	200
S2	70	81	75	95	110	105	100

We now normalize both S1 and S2 to $1 on 1 Jan 2020

- Normalization factor for S1 is 200 and normalization factor for S2 is 100

PAIR IDENTIFICATION: DISTANCE MEASURE ALGORITHM (GATEV '06) WORKED EXAMPLE

From previous slide, normalization factor for S1 is 200 and normalization factor for S2 is 100. Applying this normalization factor, we note following normalized price series from 1 Jan 2019 to 1 Jan 2020

- Divide all S1 prices by 200, and divide all S2 prices by 100

Date	1 Jan 2019	1 March 2019	1 May 2019	1 July 2019	1 Sep 2019	1 Nov 2019	1 Jan 2020
S1	0.910	0.895	0.925	0.98	1.085	0.990	1.000
S2	0.700	0.810	0.750	0.950	1.100	1.050	1.000
Spread (S1/S2) - 1	30%	10%	23%	3%	-1%	-6%	0%

For historical period 1 Jan 2019 to 1 Jan 2020 between S1 and S2, we compute historical distance between S1 and S2 as:

Define S_{mean} = average spread

Define *historical distance* = sum of squared deviations over % spread

We will now refer to the above measure as *historical distance* between S1 and S2

PAIR IDENTIFICATION: DISTANCE MEASURE ALGORITHM (GATEV '06) WORKED EXAMPLE

For each stock S_i in our S&P 100 universe, assign its "matching partner S_i" as stock with smallest *historical distance* out of all possible pairs

Assuming there is no missing data or illiquid stocks, we should now have 100 pairs

Sort all pairs in ascending order of their *historical distance*

Chose top-N pairs (with lowest *historical distance*) from our previous result to form traded portfolio.

In practice, (e.g. in Gatev '06) paper, N = 20. However, optimal value of N might be a question for backtesting to answer

Say we chose N = 20.

PAIR IDENTIFICATION: DISTANCE MEASURE ALGORITHM (GATEV '06) WORKED EXAMPLE

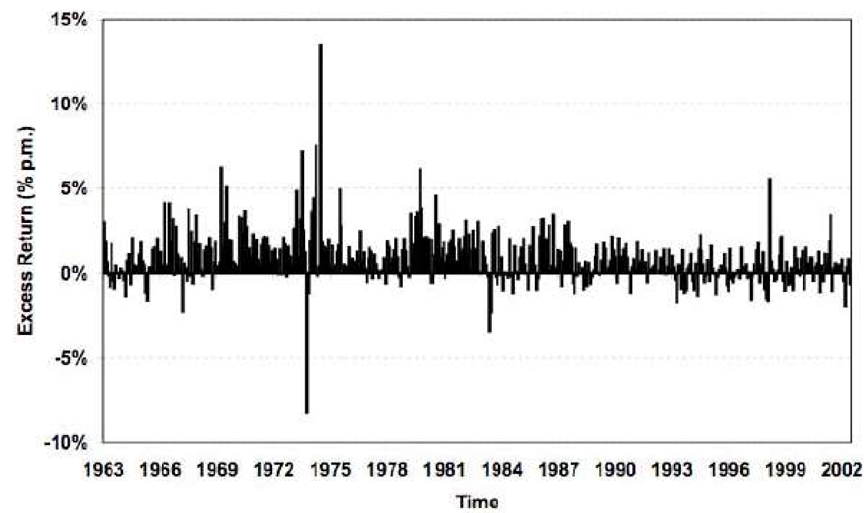
For each of 20 pairs that we chose to trade on 1 Jan 2020, we will follow this process for next 1 year (from 1 Jan 2020 to 1 Jan 2021):

When current spread (between normalized prices using normalization factor calculated on 1 Jan 2020) diverges by 2 or more historical s.d. from historical average (computed between 1 Jan 2019 and 1 Jan 2020):

- We will simultaneously open a long and short position in pair depending on the direction of divergence
- Both legs are closed (reversed) once the spread converges to 0 again
- The ratio of each stock to long and short depends on their respective normalization constants

Choice of 2 s.d. above is a calibration parameter and might depend on backtesting

Monthly excess returns of Top 20 pairs portfolio
May 1963 - December 2002



CAPITAL ALLOCATION

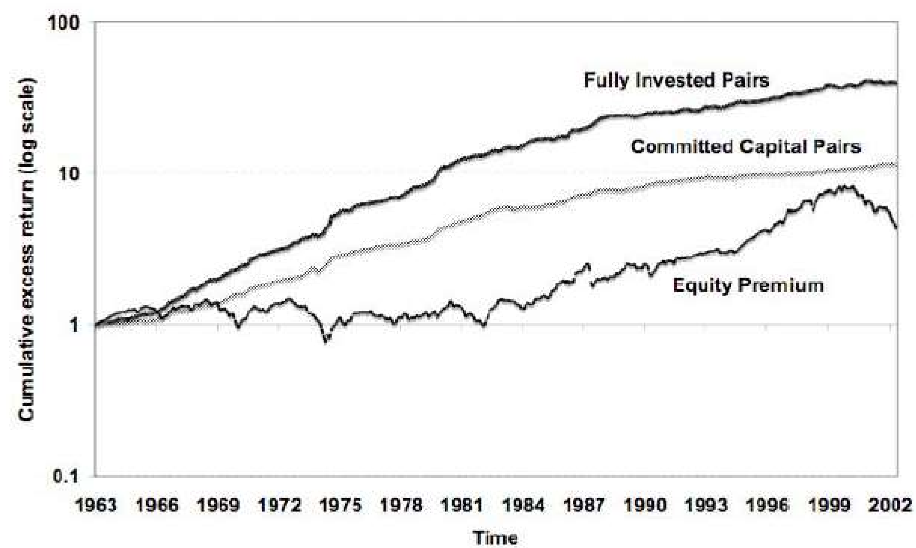
It is possible that this trading strategy will be holding a significant amount of cash and few / no equity positions from time to time

This may happen if none of the pairs breach the “2 standard deviation” trading limits

As a result, we may chose to allocate more capital to individual pairs when we are holding more cash, and vice versa

An alternative approach is to allocate capital to pairs when the deviation from historical average spread is greater

Cumulative excess return of Top 20 pairs and S&P500
May 1963 - December 2002



PAIR BREAK AND BANKRUPTCY RISK

The main risk in pair trading is that of pairs "breaking"

For instance, it is possible that leg which we are long may be for a stock that goes bankrupt

In the strategy, some form of stop loss may been required if the spread continues to widen greatly after initial entry

One way to greatly reduce bankruptcy risk is to implement pairs trading via ETFs instead of stocks

QUESTIONS TO BACKTEST

The following parameters remain to be calibrated by backtesting:

1. Stop loss
2. Capital allocation methodology between different viable pairs versus how much cash we keep
3. 2 s.d. entry parameter
4. Exit condition

CLUSTERING TOOLS

Medium Frequency Equities

PAIR RESTRICTIONS

Our prior analysis has taken the form:

i. Examine all possible pairs in our trading universe

ii. For pairs that pass some statistical criteria, we monitor them, and enter / exit each pair as actual traded prices deviate from a benchmark

Thus far, the only restriction we have placed on step (i) is to consider limiting pairs to be from similar industries

Can we refine this further?

FUNDAMENTAL CRITERIA IN PAIR RESTRICTIONS

We can also limit pair consideration along other dimensions, or the intersection of multiple dimensions:

1. Market capitalization
2. Age of company
3. Corporate finance performance metrics (together with above metrics):
   - Profitability, gross margin, cash flow from operations
4. etc

CLUSTERING IN HIGH DIMENSIONALITY

We note that intersection of each of the dimensions discussed forms a cluster

For instance, “mid cap” stocks which “recently went public” and “are profitable” might be one cluster that we can build from our first 3 dimensions. We can now rephrase our pair trading methodology as prelimiting pairs to be within the same cluster

We note that as the number of dimensions increases, the cardinality of each cluster will decrease.

Additionally, how can we:

1. Determine which dimensions are more “important” in clustering our data?
2. Construct dimensions which are “unique” from each other?

PAIRS TRADING IN SUMMARY

Pairs trading is an example of a market neutral strategy

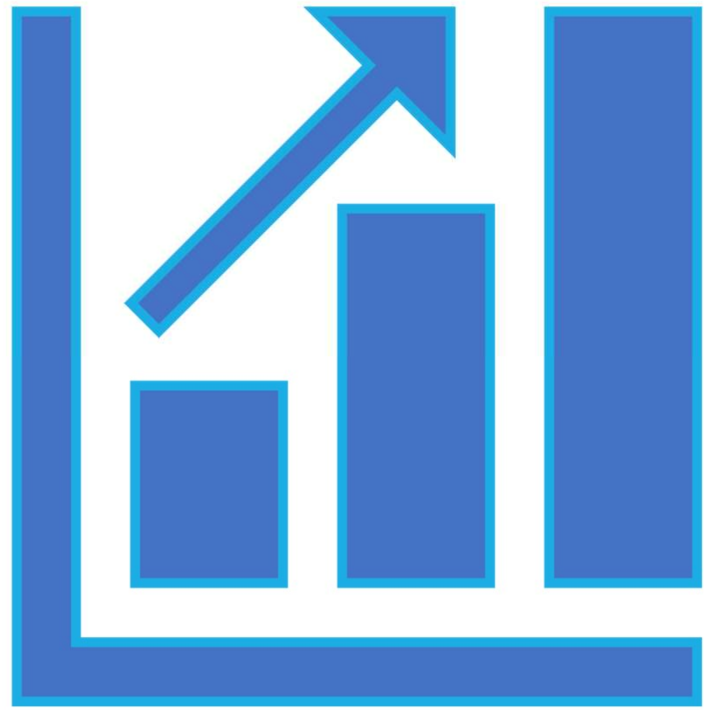
Performance is dependent on pair filtering, selection, capital allocation and entry / exit calibration, rather than overall market performance

We will examine the concept of market neutrality in more detail for the rest of this class

SUMMARIZING

In this class, we have:

- Discussed the pairs trading methodology via constructing a distance measure, correlation, as well as cointegration
- Considered how pairs trading is a specific example of a dollar neutral trading strategy
- Discussed how to transform the output of any quantitative trading strategy into both a dollar as well as industry neutral format



IN OUR NEXT CLASS:

We will evaluate commonly used time series heuristics from technical analysis

- Specifically, we will discuss use of kernel estimators from below mentioned paper to perform automated computer recognition of visual patterns:
- Lo, A.W., Mamaysky, H. and Wang, J. (2000), Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation. The Journal of Finance, 55: 1705-1765.

We will also replicate in python pairs trading, and also support vector machines for TS classification

- Code for the trading strategy will be uploaded to the class folder
- Feel free to reach out via email if you have any difficulties getting Python to install