Big Data to trade bonds/FX & Python demo on FX intraday vol

Saeed Amen, Quantitative Strategist

Managing Director & Co-founder of The Thalesians

@thalesians / commentary around finance

www.thalesians.com

saeed@thalesians.com

June 2015

Outline

- Big Data to trade bonds/FX
  - What is Big Data? What type of data is available?
  - Understanding difference between structured and unstructured news data
  - Case study: using RavenPack Macro news analytics to trade FX and bond futures
    - Discussion of RavenPack's news analytics dataset
    - Create longer term news based economic sentiment indices to mimic economic surprise indices to systematically trade FX and bond futures & applying a news volume based indicator for FX carry
- Python demo
  - Deciphering FX intraday volatility patterns

What is Big Data? What type of data is available?

What is Big Data?

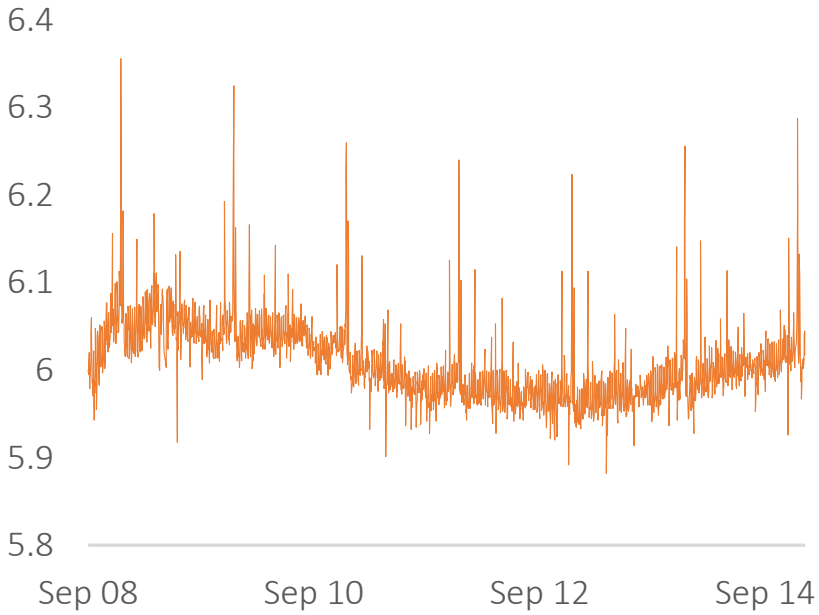
- IBM: Every day, we create 2.5 quintillion bytes of data — so much that 90% of the data in the world today has been created in the last two years alone. This data comes from everywhere: sensors used to gather climate information, posts to social media sites, digital pictures and videos, purchase transaction records, and cell phone GPS signals to name a few. This data is **Big Data**.
- Wikipedia: **Big Data** is an all-encompassing term for any collection of data sets so large or complex that it becomes difficult to process them using traditional data processing applications.
- But how many people really use Big Data, in particular in trading?
- Not as many people as you might think!

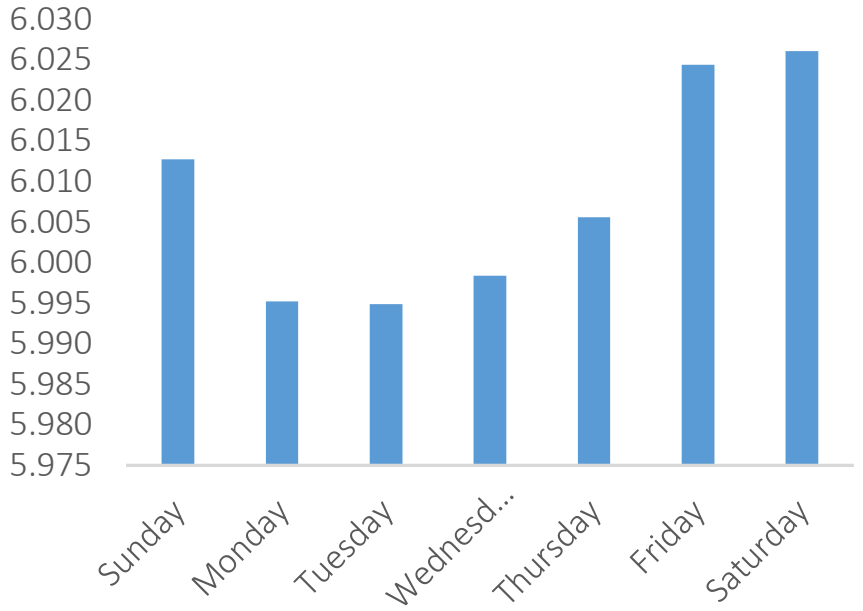
What type of data is available?

- News wires – Dow Jones, Bloomberg, Reuters
- Social media – Twitter, Flickr, Facebook
- Web – Wikipedia, blogs, PR news releases
- Search – Google
- Examples of large structured data sets
  - Bloomberg – terminal product and machine readable
  - Reuters – machine readable news
  - RavenPack – Dow Jones, web and PR machine readable
  - Knowsis – Twitter – web product and machine readable
  - Estimize – crowdsourcing analyst estimates – web product and machine readable
  - And many more…

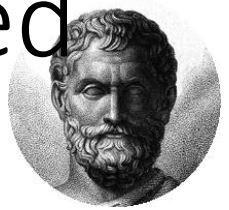
A fun example

- Happiness measured via tweets – Hedonometer Index from University of Vermont – we can actually extract trading signals from this!





Understanding difference between structured & unstructured news data



- Unstructured news data
  - Read news articles, blogs etc. in their raw text form and then directly apply text based analysis to gauge sentiment
  - Need to handle large amounts of data and also need to do natural language processing, which is non trivial
- Structured news data
  - Vendors processes a large amount of news from numerous sources into a more manageable dataset for us to explore
  - Data more easily accessible with precomputed sentiment scores/volume reading
  - Traders can concentrate on creating effective trading rules and running risk, rather than spending that time dealing with massive quantities of unstructured news and text analysis.

Case study: using RavenPack Macro news analytics to trade FX and bond futures

Automating news filtering

- Using news to trade markets is not new idea
- A trader essentially “filters” news into the “signal and the noise”
- How can we read news in automated fashion
  - Raw news data stream - apply text analysis on actual news articles
  - Use pre-made news time series generated by vendors using keywords/topic identification
- Easier to use pre-made news time series
- However, what news filters do we use?
- How can we convert news time series to buy/sell?

General approach

- Steps to follow to understand how news data can be used to trade markets
  - First: Analysing text from news articles/headlines (RavenPack does this step!)
  - Second: Aggregate RavenPack data by time frequency
  - Third: Create a sentiment index for specific areas of interest
  - Fourth: Apply a trading rule to the sentiment index

First: Analysing text from news articles/headlines

- Use structured news data, which reduces our workload
  - RavenPack Macro 4.0 – Web, Dow Jones and PR editions combined
- For each news event captured a record containing a number of different fields are output
  - Timestamp of publication – In UTC time with a millisecond timestamp
  - Focus of the publication – Includes details on the country and the general subject of the news
    - Use these fields later to filter news for example for US news related to the economy
  - Positive/negative nature of news – Scaled from 0 to 100, where >50 is positive, <50 is negative (and 50 is neutral)
    - Use this later for identifying the bullishness/bearishness of an article for trading purposes (Event sentiment score – ESS)
  - Measures of the relative novelty of news – "Newer" news as opposed to repeated headlines scores higher (Event Novelty score – ENS)
  - Prevalence of news – Identify the number of positive or negative events for a certain entity (Aggregate Event Sentiment – AES) and also the general news volume on an entity (Aggregate Event Volume – AEV)
  - Source of the news and the RavenPack product edition
  - A full list is available from RavenPack

Second: Aggregate RavenPack data by time frequency

- Our objective is to create a daily signals trading at local bond market closes (differs for each bond market)
- RavenPack data is available throughout the day on a high frequency basis
- Gives us control over a precise cut off point for trading purposes

Third: Create a sentiment index for specific areas of interest

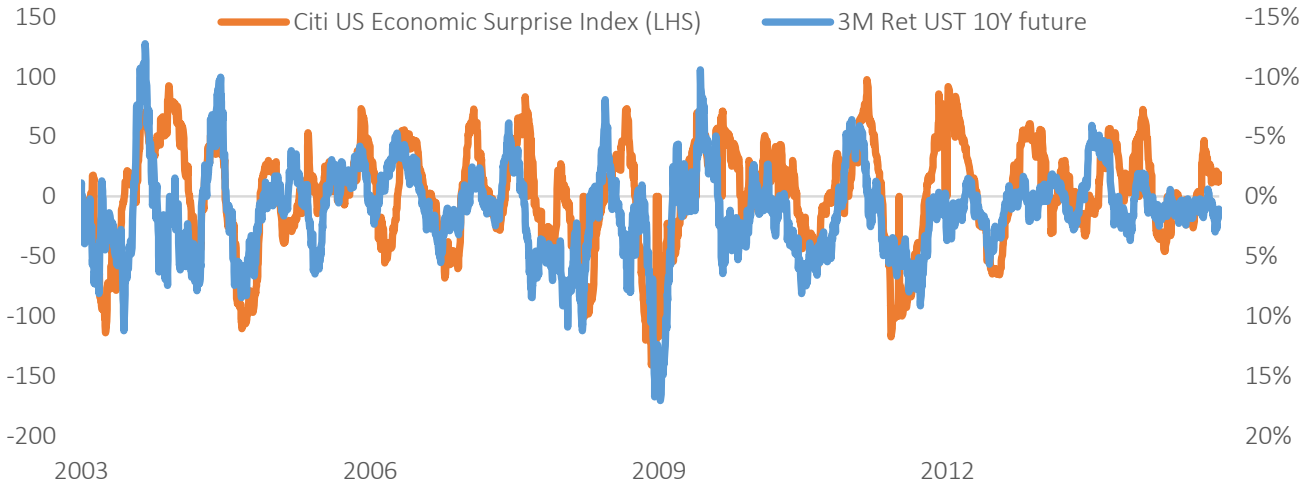
- Topics are
  - business
  - economy
  - environment
  - politics
  - society
  - Use as an input ESS – to give positive/negative signals
- Topics are further subdivided into groups
- We shall mainly focus on the “economy” topic

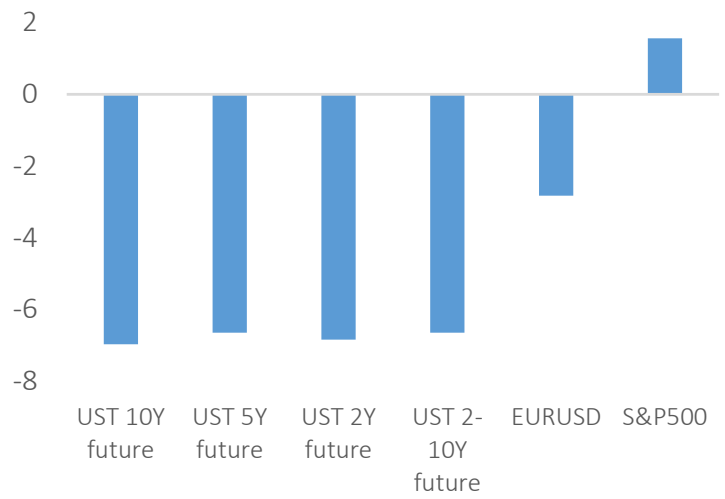
Fourth: Apply a trading rule to the sentiment index

- Apply our sentiment filter to trade assets
- Our focus will be on G4 sovereign bonds (US, Japan, Germany and UK)
- We have assumed equal sized notionals for our bond futures baskets and also reported returns in local currency
- In practice, you might consider vol or liquidity weighting bond exposure and report in an investors home currency (as well as considering hedging)
- Later, small section on FX

Economic surprise indices

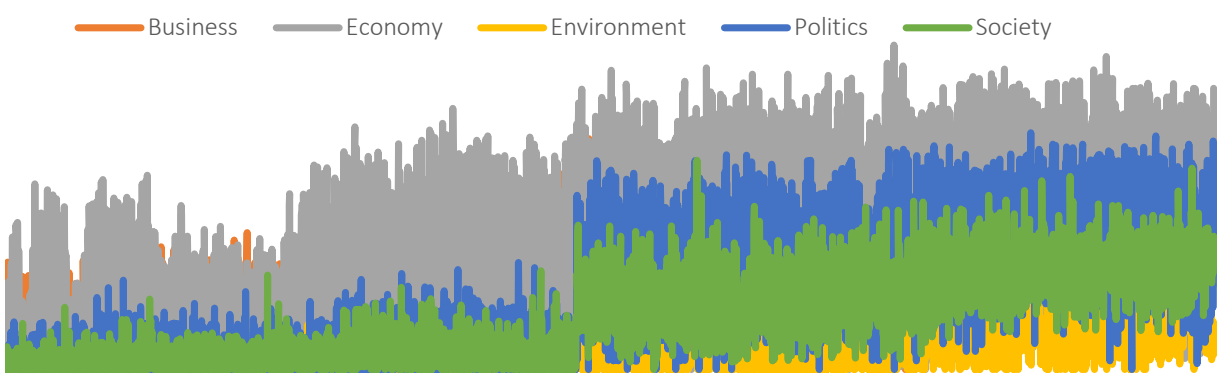
- Understand the relationship between economic surprise indices and macro assets
- Clearly, there is a strong correlation, we can also fade longer term moves
- Can we mimic using news data, which corresponds to a richer set of events?

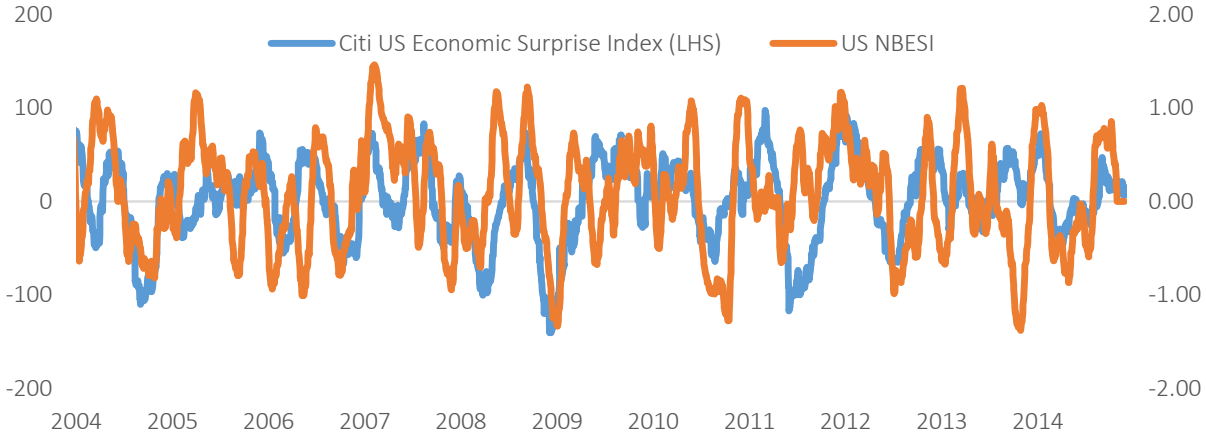




News based economic sentiment indices

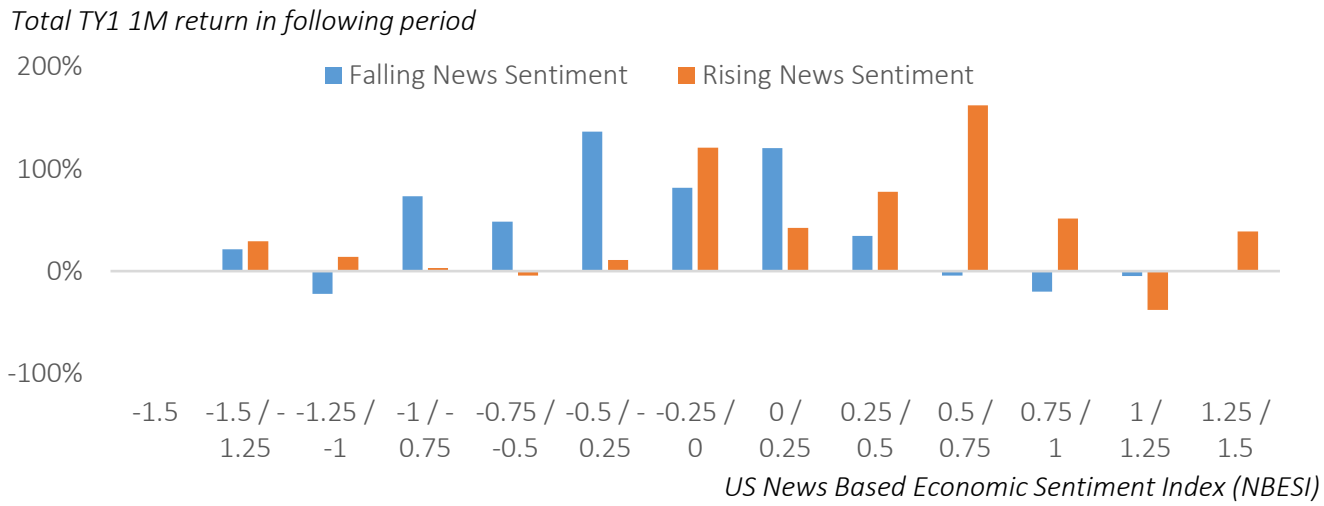
- Use the “economy” topic classified by RavenPack
- Aggregate high frequency news data into daily data
- Apply smoothing and adjust for volume changes
- Objective is to create slower moving sentiment indicator





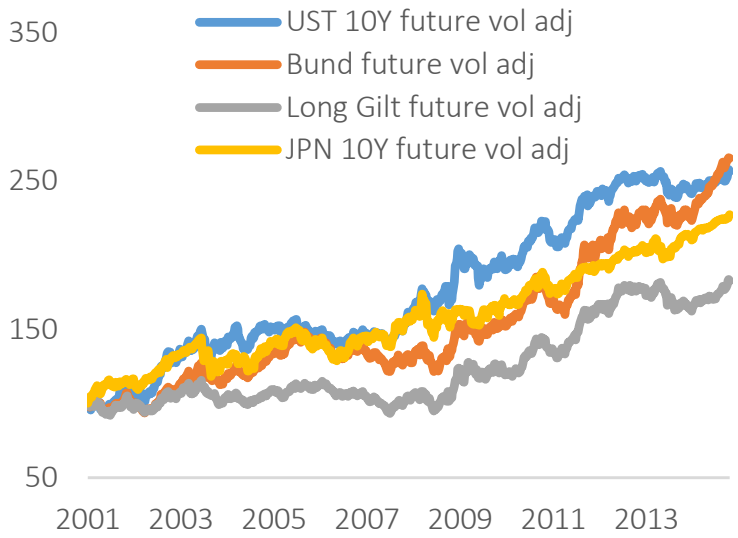
Distribution of returns based on NBESI

- We have examined what returns are for UST 10Y futures for an entire month
- We have bucketed the values for falling/rising news sentiment
- Large numbers, because we over 15 years and we enter a new position every day



Relationship between G4 bond markets

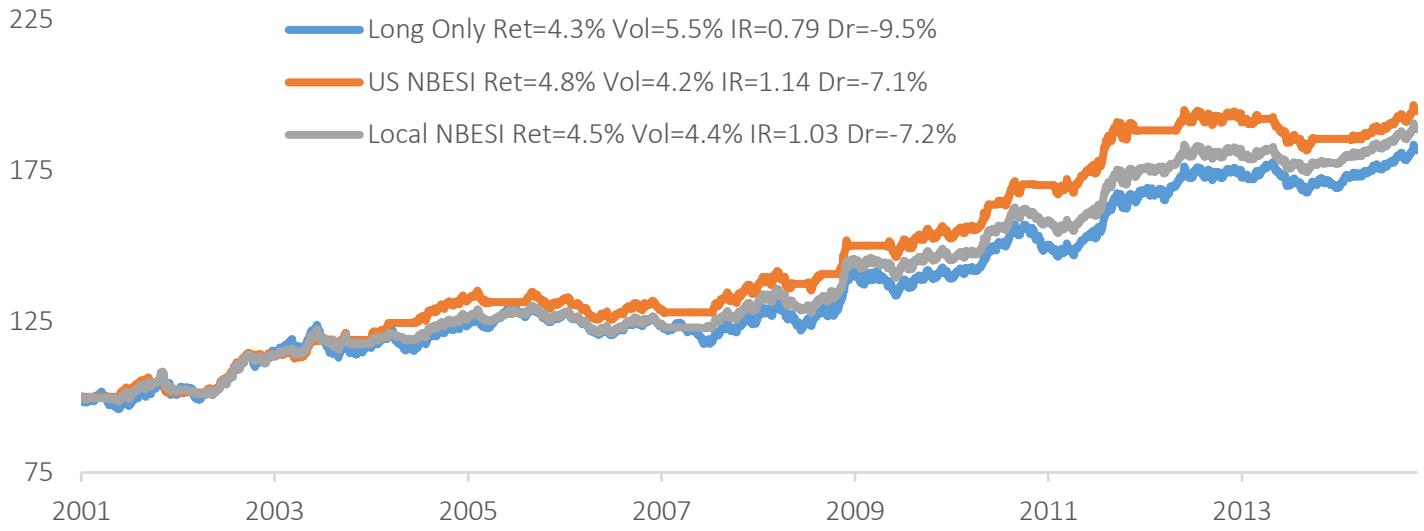
- Vol adjust G4 returns – tend to be highly correlated
- We shall utilise this observation to also use US NBESI as filter for rest of G4
- Rationale is that US is a major driver for other bond markets

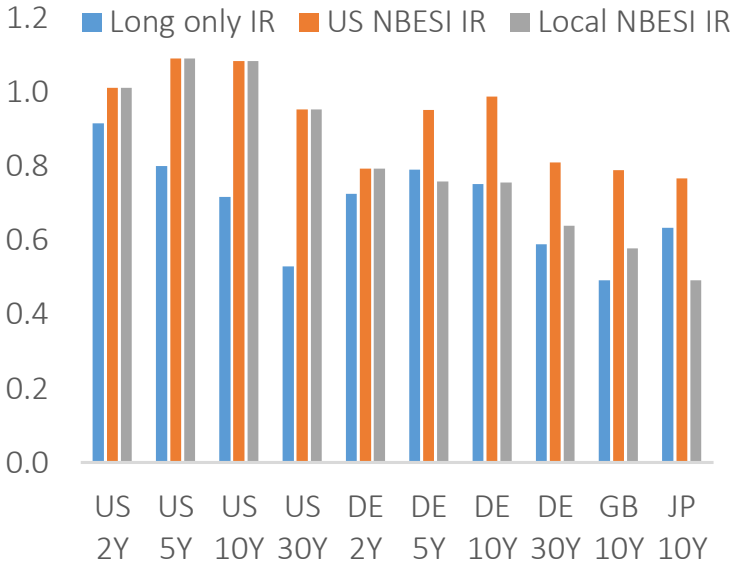


	US10Y	DE10Y	GB10Y	JP10Y
US10Y		68%	65%	35%
DE10Y	68%		85%	38%
GB10Y	65%	85%		33%
JP10Y	35%	38%	33%	

Trading G4 bond futures with NBESI

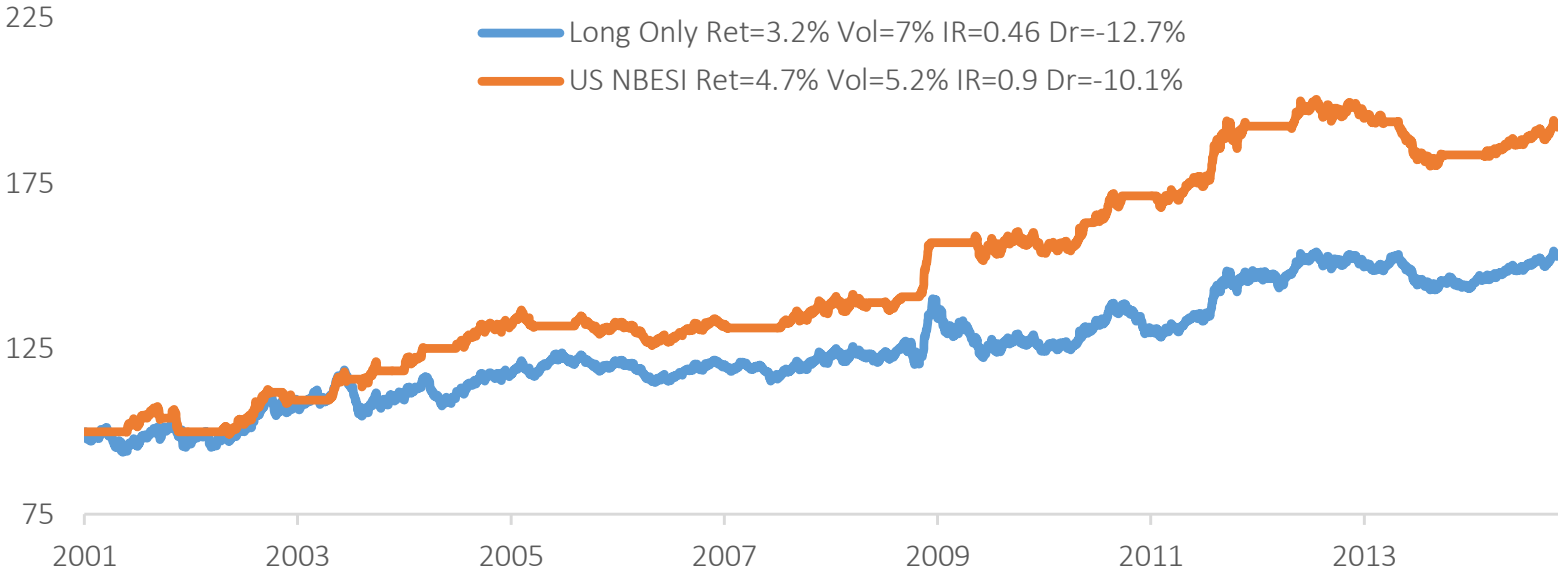
- Buy bonds when NBESI is falling and negative (ie. worsening economic environment which is typically associated with better bond performance) and hold position till we reach an extreme
- Filtering by US NBESI has best returns, followed by local NBESI
- Includes transaction costs & we have adjusted futures to create a continuous series (roll costs executed)

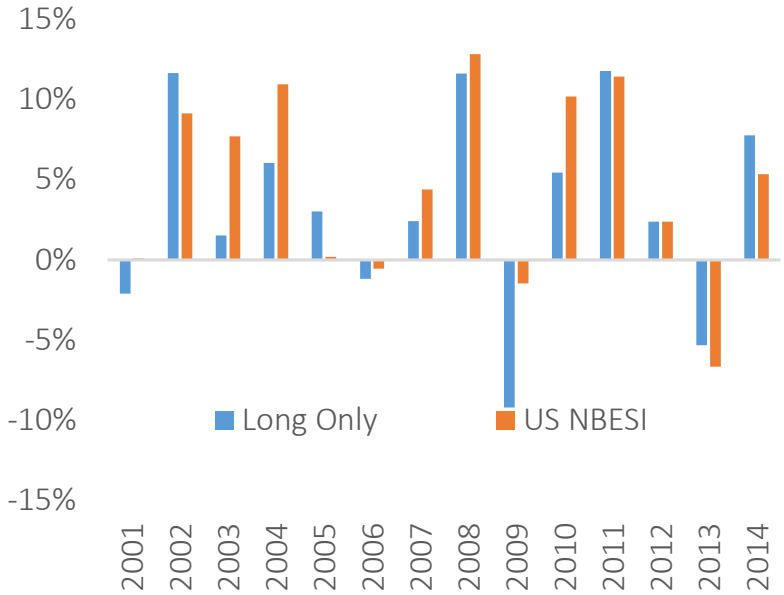




Trading US bond spreads with NBESI

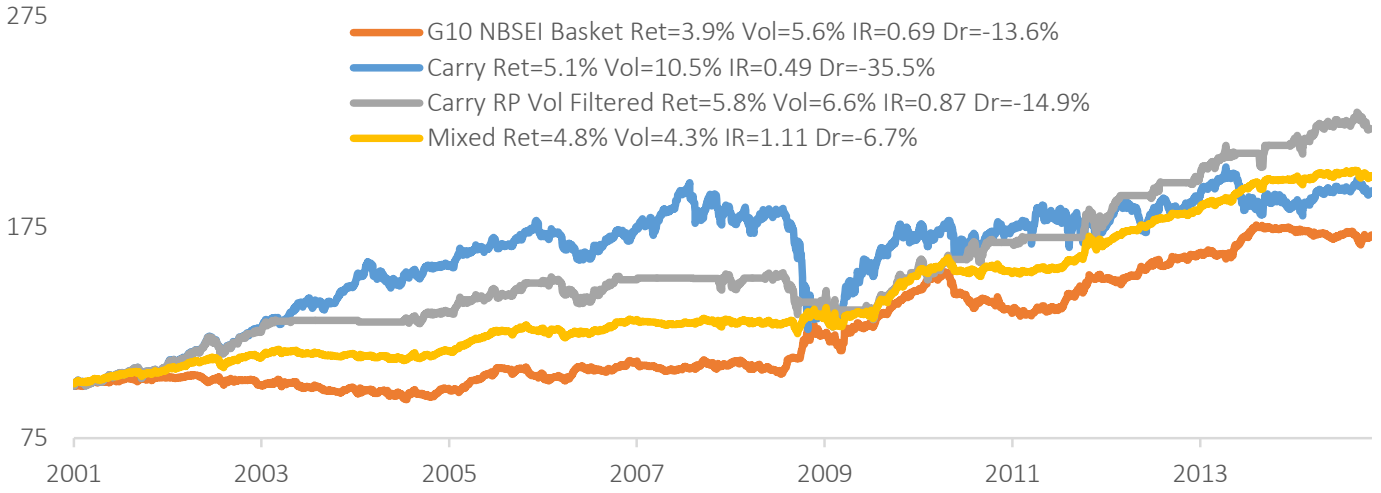
- Applying a similar trading rule we trade spreads across the US Treasury curve
- Buy longer dated bond futures and fund with shorter dated ie. a carry style trade
- Applying US NBESI trading rule considerably outperforms passive strategy

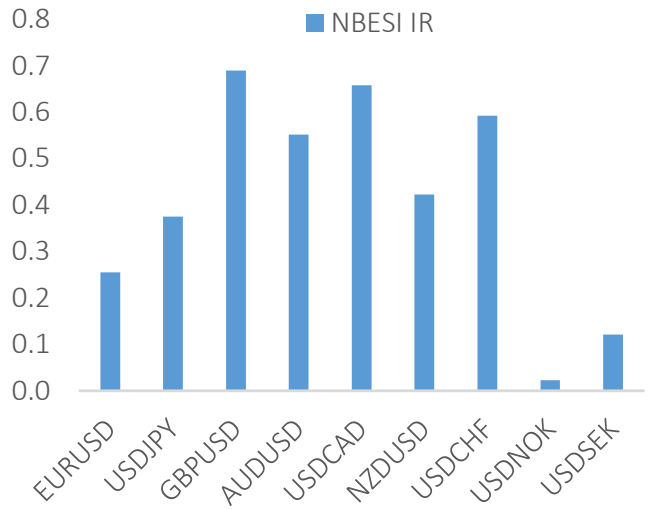




Trading FX with NBESI

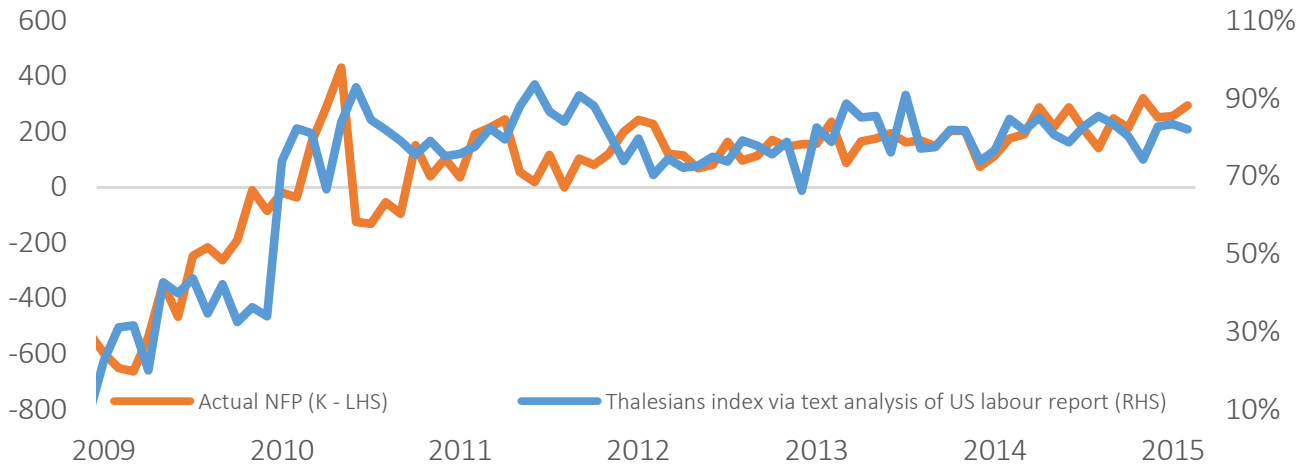
- Fade moves in NBESI, ie. falling NBESI => buy, rising NBESI => sell
- G10 FX NBESI basket combined with filtered G10 FX carry strategy (using RavenPack news volume) has better risk adjusted returns than either
- Carry filter used a different aggregation scheme to NBESI





One last thing on news…

- Using unstructured data
- Creating an index of US labour market – work in progress!



Python Demo: What will we use?

- A lot of analysis for Big Data happens in Python
- Built PyThalesians coding library
- Used following libraries heavily
  - NumPy
  - SciPy
  - Pandas
- Strongly recommend using Python!
- Will be deciphering intraday FX volatility patterns!

What will we do?

- Market data – load using live Bloomberg API call and cache (retail FX data), then plot
- Analyse intraday vol – calculate, then plot
- Contrast tick count & intraday vol, then plot
- Focus on scheduled data events, calculate, then plot
  - spot moves
  - surprise vs. spot moves
  - intraday vol
  - plot

Conclusion

- Examined using news data to trade
  - Using RavenPack data
  - Trading strategies for bonds and FX
- Used high frequency data to analyse FX market
  - Discovered exciting patterns in intraday vol
  - Examined how price action behaved around data events

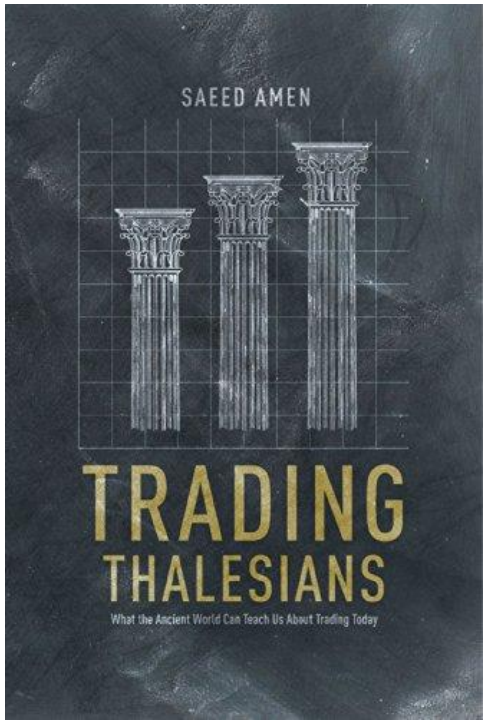
If you liked my talk…

- Any questions, ask now or e-mail at saeed@thalesians.com or tweet @thalesians
- Read my book ***Trading Thalesians – What the ancient world can teach us about trading today***
- The Thalesians publish quant strategy papers (short versions free)
- The Thalesians consultancy
  - Create systematic trading models
  - Bespoke financial market research based projects
  - Particular expertise in FX (you probably guessed!)
- **And yes, I trade my own money – skin in the game!**

…leave you with some words from my book

A yard to end

Beneath the whisper are words to rise and strive,
That seek to break apart the sentiments,
Which say no and never and don't and won't,
To trounce the dreams, to pounce upon what seems,
Go wipe away the false and grasp the truth,
So seek the path and endeavor and rise,
Go forth traverse the lows and climb the highs,
Ignore the pain, the fails and the near just,
A yard to end, that goal it'll be your pole,
And turn to back and smile in that last mile,
Year next will rise in mind as you'll sure find,
If simple were life, boredom would be rife!



www.thalesians.com

THE THALESIANS

Θαλῆς ὁ Μιλήσιος

Company

In 2009 / Established by Imperial College graduates with investment banking experience
Hosted / Over 100 seminars internationally in London, New York, San Francisco & Budapest
Speakers / From quant finance world & journalists such as Greg Zuckerman of WSJ
Now / Introducing quant consulting and research services

Experience

Decade / Creating & running systematic trading models at Lehman Brothers, Nomura & Thalesians
Co-developer / Lehman's MarQCuS systematic trading model with $2bn AUM
Author / *Trading Thalesians: What the ancient world can teach traders today**

Consulting

Bespoke / Systematic trading models for alpha generation
Expertise / Quant analysis, FX hedging, TCA & much more

Research

Subscribe / Ground breaking publications describing innovative approaches to quant trading
Includes / Work on 4pm FX fix featured in WSJ & Google/Bloomberg News data to trade markets

Saeed Amen
Managing Director & Co-founder
Quantitative Strategy
p // +44 20 3290 9624
e // saeed@thalesians.com
t // @thalesians
w// www.thalesians.com

* preliminary title & expected publication on Palgrave Macmillan in October 2014

Paul Bilokon, Saeed Amen & Matthew Dixon
Founders of THE THALESIANS

The Thalesians are a think tank of dedicated professionals with an interest in quantitative finance, economics, mathematics, physics and computer science, not necessarily in that order.

www.thalesians.com