BANK NEGARA MALAYSIA
CENTRAL BANK OF MALAYSIA

EUROPEAN CENTRAL BANK
EUROSYSTEM

Per Nymand-Andersen
Adviser, European Central Bank

Technology transformation in central banking

Big Data: Strategies and Advanced Analytics

Monetary Policy 2.0? Conference 2017

Bank Negara Malaysia, 24 July 2017

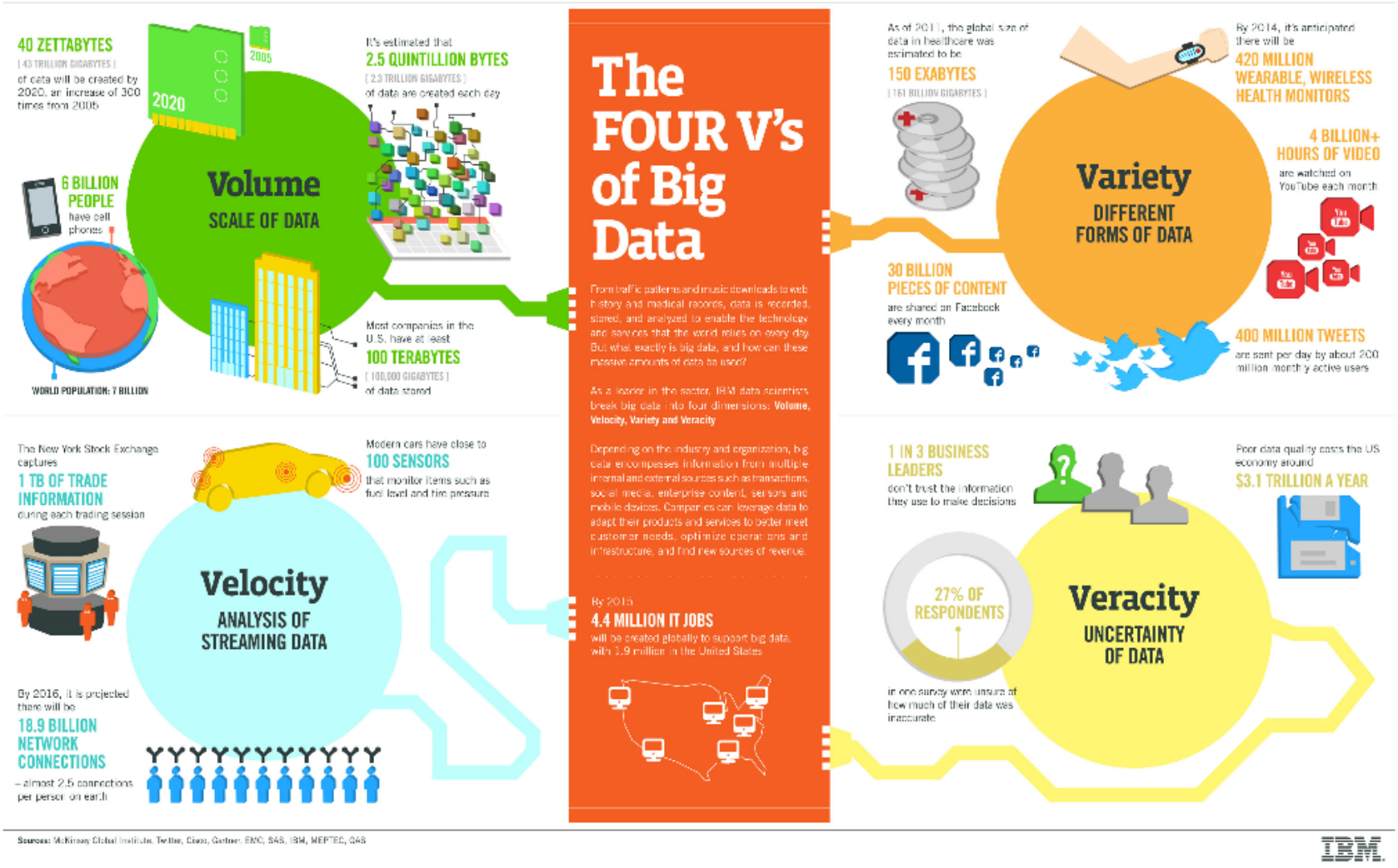
"Progress lies not in enhancing what is, but in advancing towards what will be" (Khalin Gibran)

Overview

1. **Pretty Big Data - Reflections for policy purposes**
2. **Embracing technology and Advanced Analytics**
3. **Discovery, collaboration and Partnership**

Ref.: **"Big data: The hunt for timely insights and decision certainty - Central banking reflections on the use of big data for policy purposes**, IFC working Paper No 14, 2016, Per Nymand-Andersen

Pretty Big Data - The "4V"

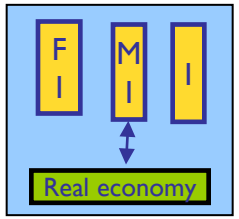


Source: IBM: http://www.ibmbigdatahub.com/sites/default/files/infographic_file/4-Vs-of-big-data.jpg

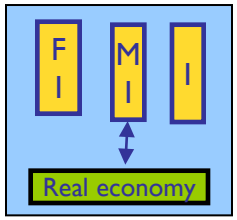
Data sources for Monetary Policy & Financial Stability

Other systems ⟷ **Financial system** ⟷ **Real economy**

Financial systems



Financial systems



Other sources

Policy events,
fiscal policies,
Global economy
External chocks
Sector failures

Transmission of central banking policies, managing expectations, feed-back loops

Regulatory and oversight framework

Financial Intermediaries (FI)

Banks
Insurance, pension
hedge funds,
Rating agents, others

Markets and instruments (MI)

Securities,
Derivatives
and other products

Financial infrastructures (I)

Pre-trading,
Trading,
post-trading
Infrastructures
Clearing,
settlements

Raw sources:

Pre-trading, trading and post trading activities, actors, instruments, platforms, volumes, prices, frequencies, maturities, regulators and overseers alike

Derived indicators:

Yield curves, density, liquidity, price measures,
Functioning of markets, instruments and post trading activities
Risks measures and contagions indicators
Imbalances and concentrations measures
Effects on performance (Collective behaviour, Interconnections networks, transmission of instability measures

CORPORATE SECTOR

HOUSEHOLD SECTOR

GOVERNMENT SECTOR

FOREIGN SECTOR

Reflections on use for central banking policy purposes

Supplementary Insights

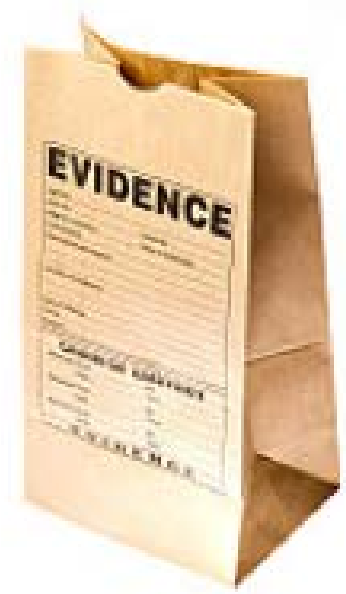
Near real-time snapshot

Early warning indicators

Detecting trends and turning points

Extract information on the impact of policy actions

New theories from combining different science fields



Reflections for central banking policy purposes

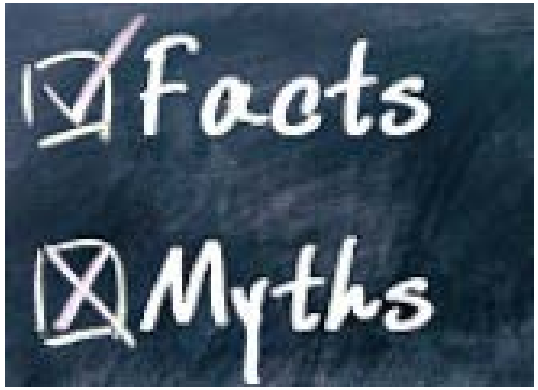
Better predictions

Means for stimulating debate

An opportunity to adjust model-based theory

More timely and frequent information

banking supervision and regulators needs micro data



Data mania versus phobia – a new paradigm of records

Digital transformation in finance and economics

E- trade

Settlement systems

Credit cards

Mobile trans

Lending & financing

- Fintech
- DLT
- D-coins
- S-contracts

Price scans

S- media

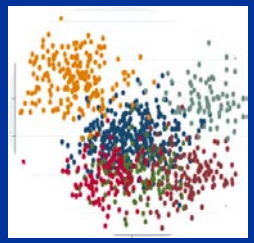
Big data

Data Scientist

Data Analytics

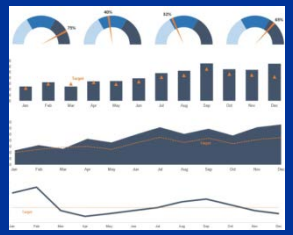
Data lab

Systematic acquire, Process, summarize



Statistical algorithm and data explorations

Packaging data for Insights & business



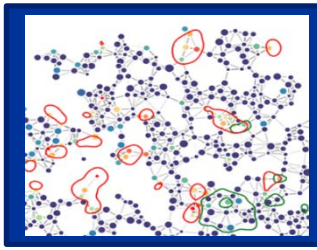
4 building blocks of Data Analytics

Digital transformation in finance and economics

Sense

Descriptive statistics analytics

Respond



Statistics pattern analytics

Predict

Now casting and forecasting analytics
Statistical modelling

Act

Prescriptive statistics analytics

Conceptual framework for showcasing big data

5 steps on the way for exploring Big Data's relevance

Input

- Source
- What info
- Coverage
- Access
- Formats
- Language
- Standards
- Free or cost

Quality

- Methodology
- Units
- Sample
- Representative
- Frequency
- Quality and updates
- Reliability
- Sustainable

Production

- Hardware
- software
- Process
- Aggregations
- Testing
- Models
- Algorithm

Results

- What can be produced
- Frequency and timelines
- Behaviour
- Purpose
- Feasibility

Assessment

- Summary
- Purpose
- Production
- + / -
- Business model
- Costs
- Benefits
- Timetable

2 remarks & 2 lessons on Big Data and Analytics

1. One misperception of big data is that we *do not need* to worry about sample bias and representativeness, as large volumes of information supersede standard sampling theory, since big data provide census-type information

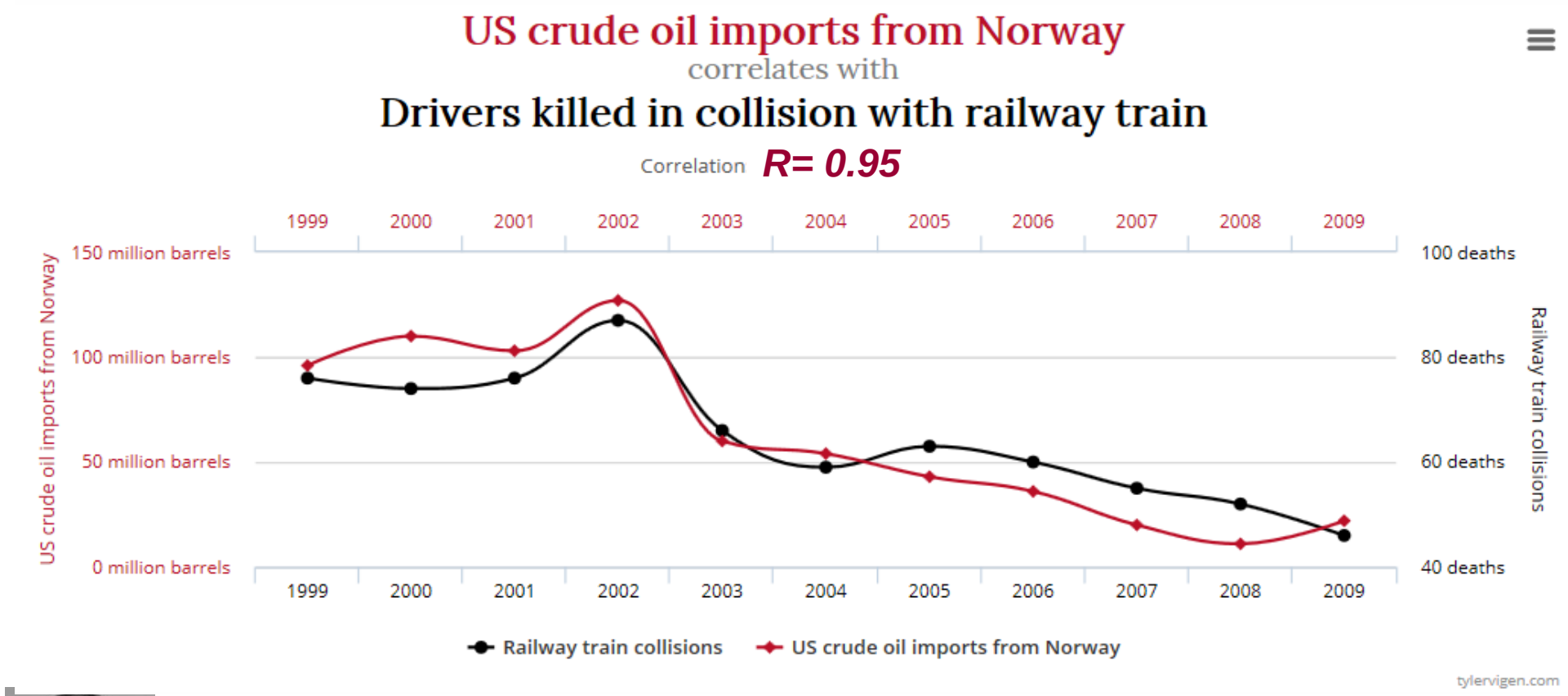
Studies done on Corporate, consumer, household analysis & indexing, text mining, pulse of the economy, consumer emotions, stock market correlations.

- Access to all tweets *means access* to the characteristics of the "tweeting" population, which may differ from those people/corporates who do not tweet
- Not all groups are represented. 21% of online adults use Twitter, varies according to age, gender, income, education, ethnic origin and country;
- Tweets vs unit measurement, double counting (tweeting and re-tweeting), over-representativeness, statistical corrections are needed
- An event driven source – volume changes do not necessary refer to reporting units nor to changes in demand

Two remarks on Big Data and Analytics

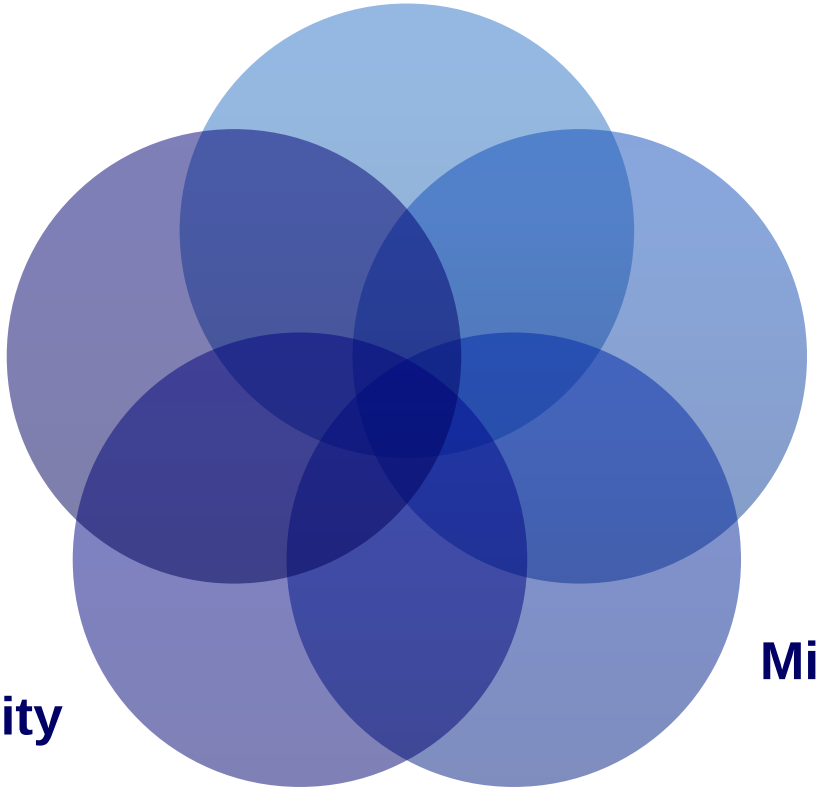
2. Correlation is *not (necessary)* causation

No conclusion can be drawn simply on the basis of correlations between two variables. The similarity is a coincidence. We say that there is no causation



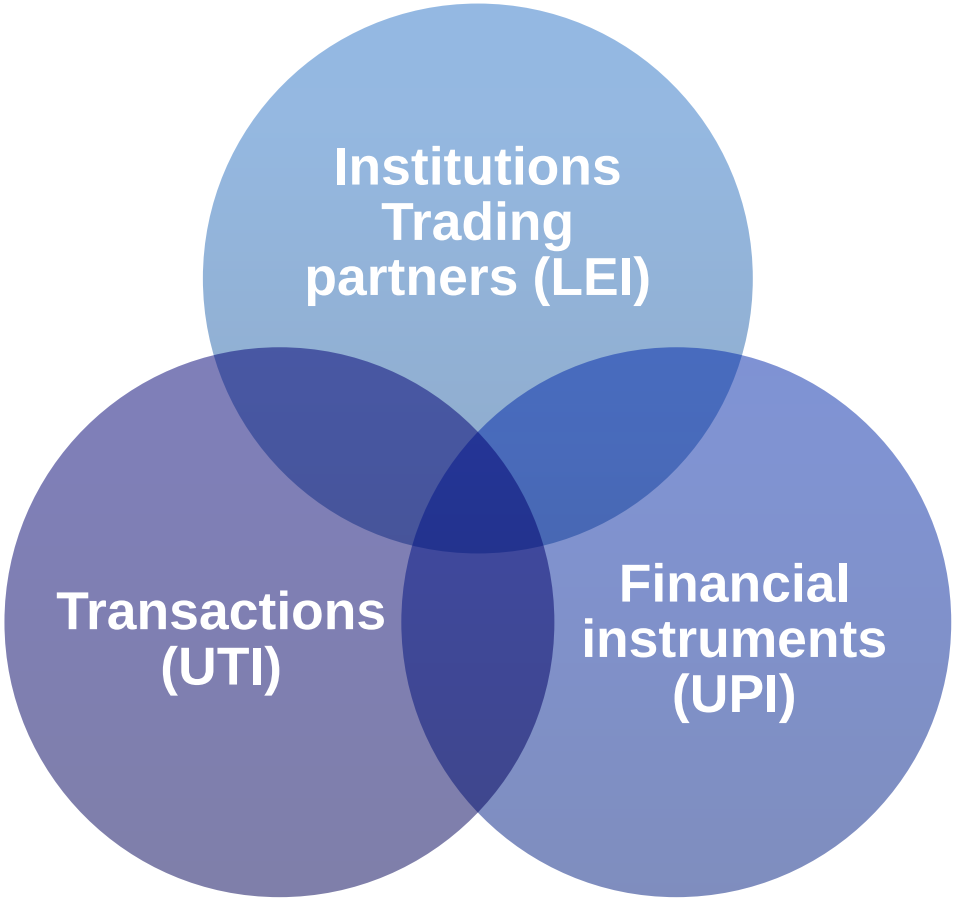
"The invalid assumption that correlation implies cause is probably among the two or three most serious and common errors of human reasoning"
Stephen Jay Gould, American evolutionary biologist and author, 1981

A source is subject to statistical quality standards



Standardisation and harmonisation ! The call for Big Data

- Big data is mainly a bi-product and not the core business of the sources
- Standardisation is a pre-condition for managing & combining large datasets
- Pre-trading, trading and post trading activities - ISO20022 and SEPA
- Money Market Reporting & Banks' loans to corporates and households



Showcasing Big data

The Irving Fisher Committee

49 central banks collaborating in showcasing Big Data

i. **Administrative dataset** (e.g. Fiscal and corporate balance sheet data)

ii. **Web search dataset** (e.g. Google type search info, web-prices)

iii. **Commercial dataset** (e.g. credit cards, operations)

iv. **Financial market dataset** (e.g. frequency trading, price and bid-offer spreads)

Preliminary results expected still in 2017

Showcasing "big data" for macroeconomic purposes

- **Source: Google search data**
- **Objective: Nowcasting euro area unemployment**

- Dataset: Google search data (google search machines)
- Data from 14 countries
- Google's taxonomy of categorising search terms, includes 26 main categories and 269 sub-categories (Finance and Banking)
- Google search data is an index of weekly volume changes
- The volumes are normalised starting at 1.00 and next week value shows the relative change of Google searches within the category (no absolute volumes)

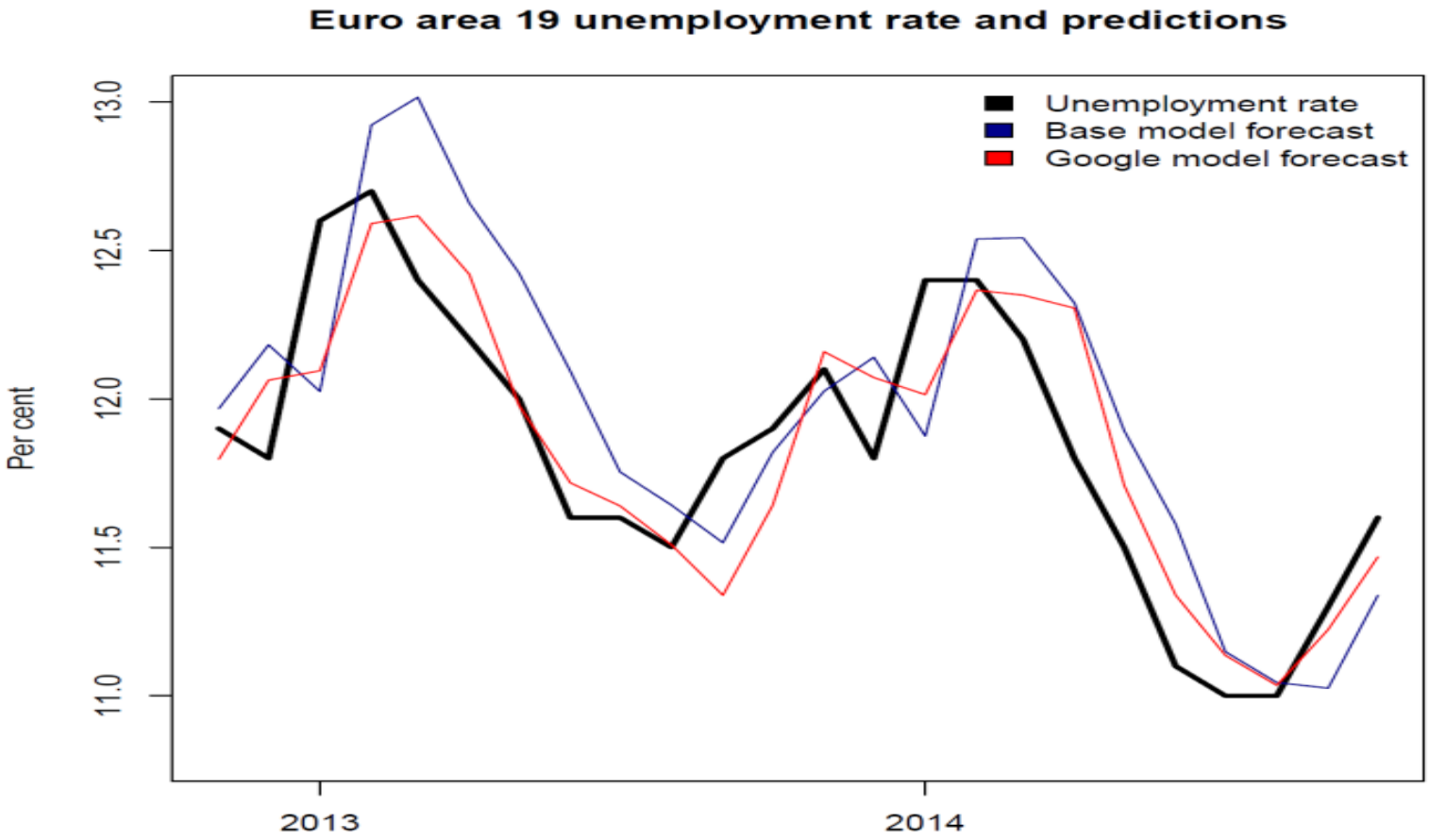
Showcasing "big data" for macroeconomic purposes

Two autoregressive models are used to nowcast euro area unemployment rate

$$log(y_t) = a + b^* log(y_{t-1}) + c^* log(y_t\text{-}_{12}) + e_t,$$

$$log(y_t) = a + b^* log(y_{t-1}) + c^* log(y_t\text{-}_{12}) + G + e_t,$$

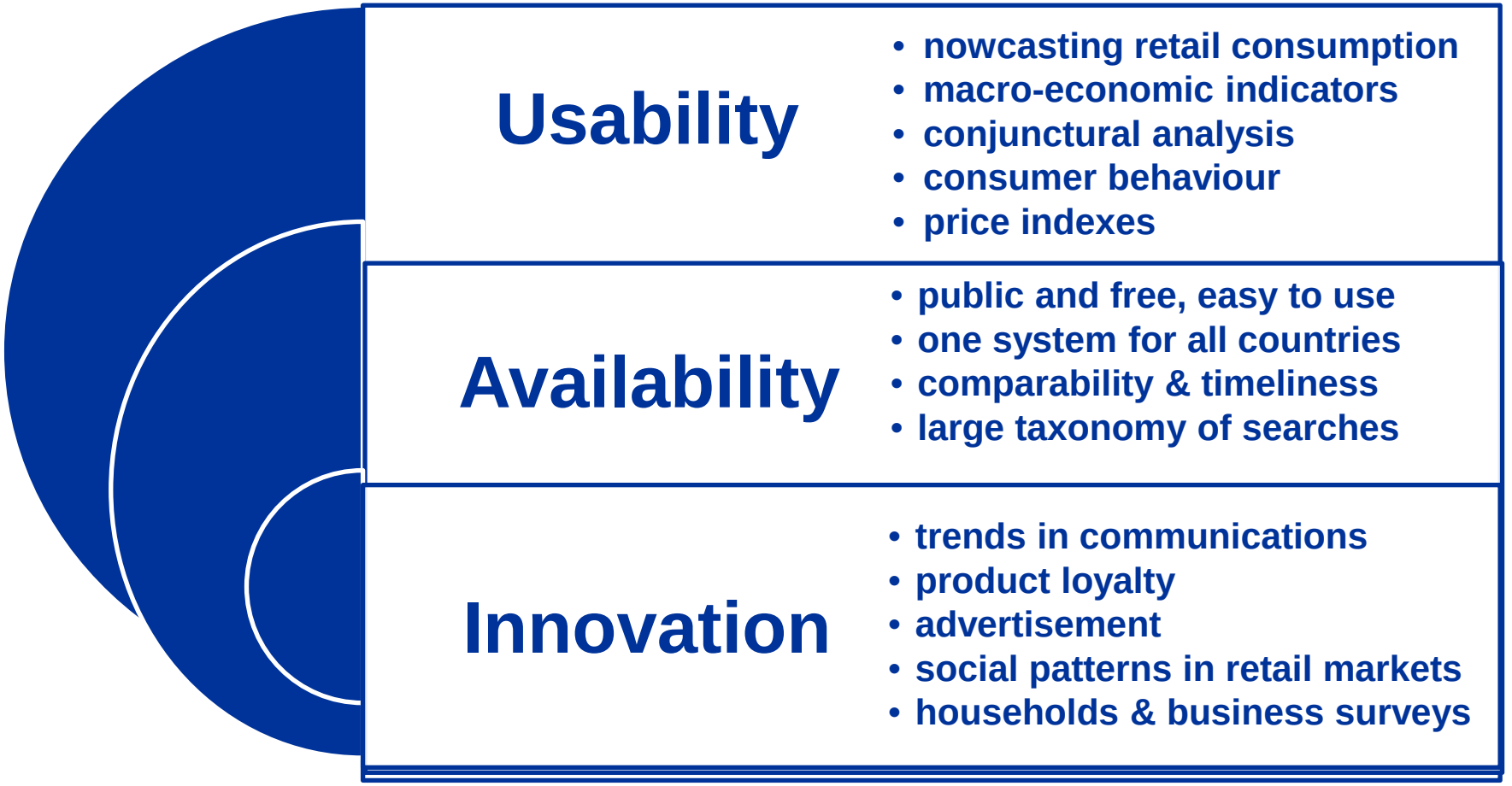
Where $Y_{(t)}$ is the unemployment rate at month($_t$) & G is the google search index



➢ ***The improvement (reduction in the errors (MAE)) range from 18.1% to 28,7%***

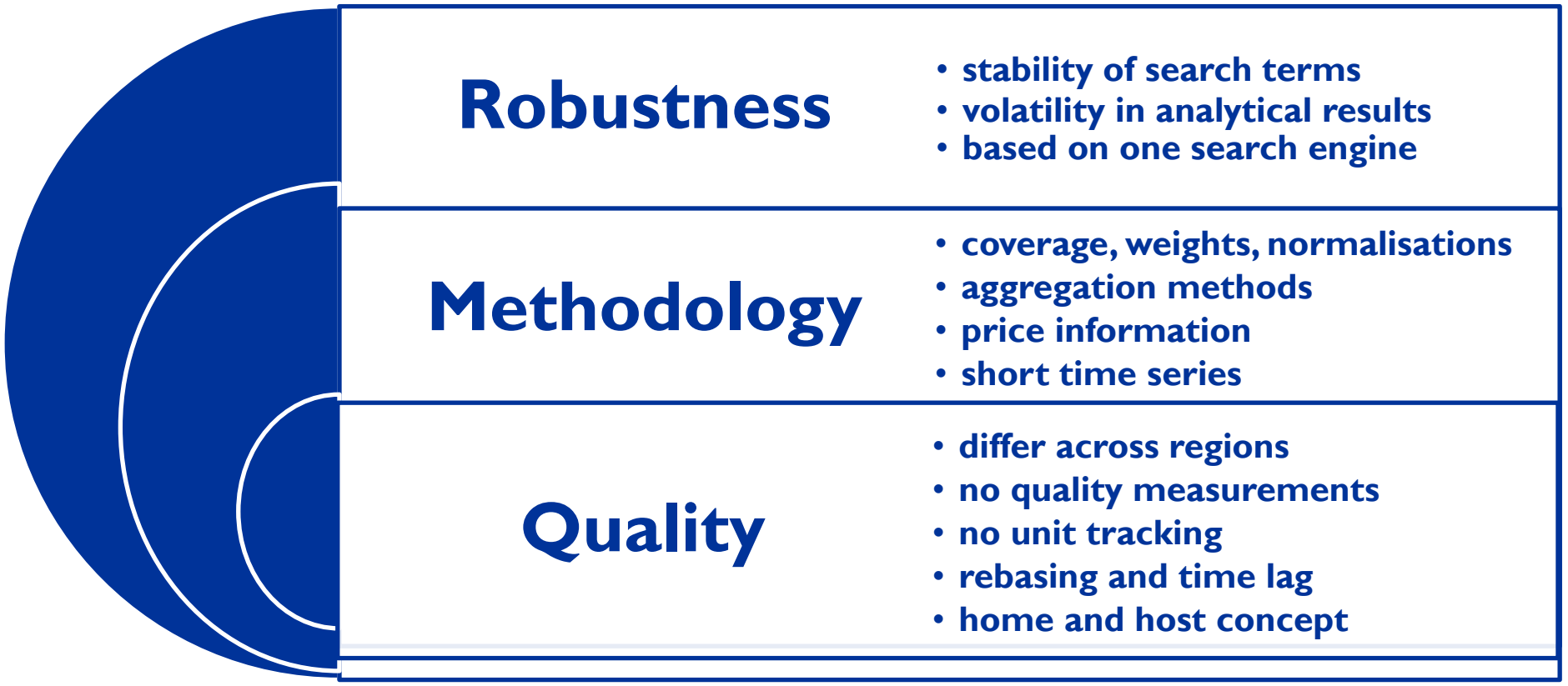
Showcasing "big data" for macroeconomic purposes

Lessons learned and way forward



Showcasing "big data" for macroeconomic purposes

Lessons learned and way forward



Take 5: Technology transformation and data analytics

1. There *is* a business case for exploring "big data" for central banking using digital sources from the financial system and beyond
2. The set-up, *competencies are diverse* and source specific
3. Use a conceptual framework & standardising financial markets
4. Join forces in private-public-university partnerships
5. Pretty Data Analytics will change behaviour, business models, knowledge & theories and regulators – Explore digital footprints !

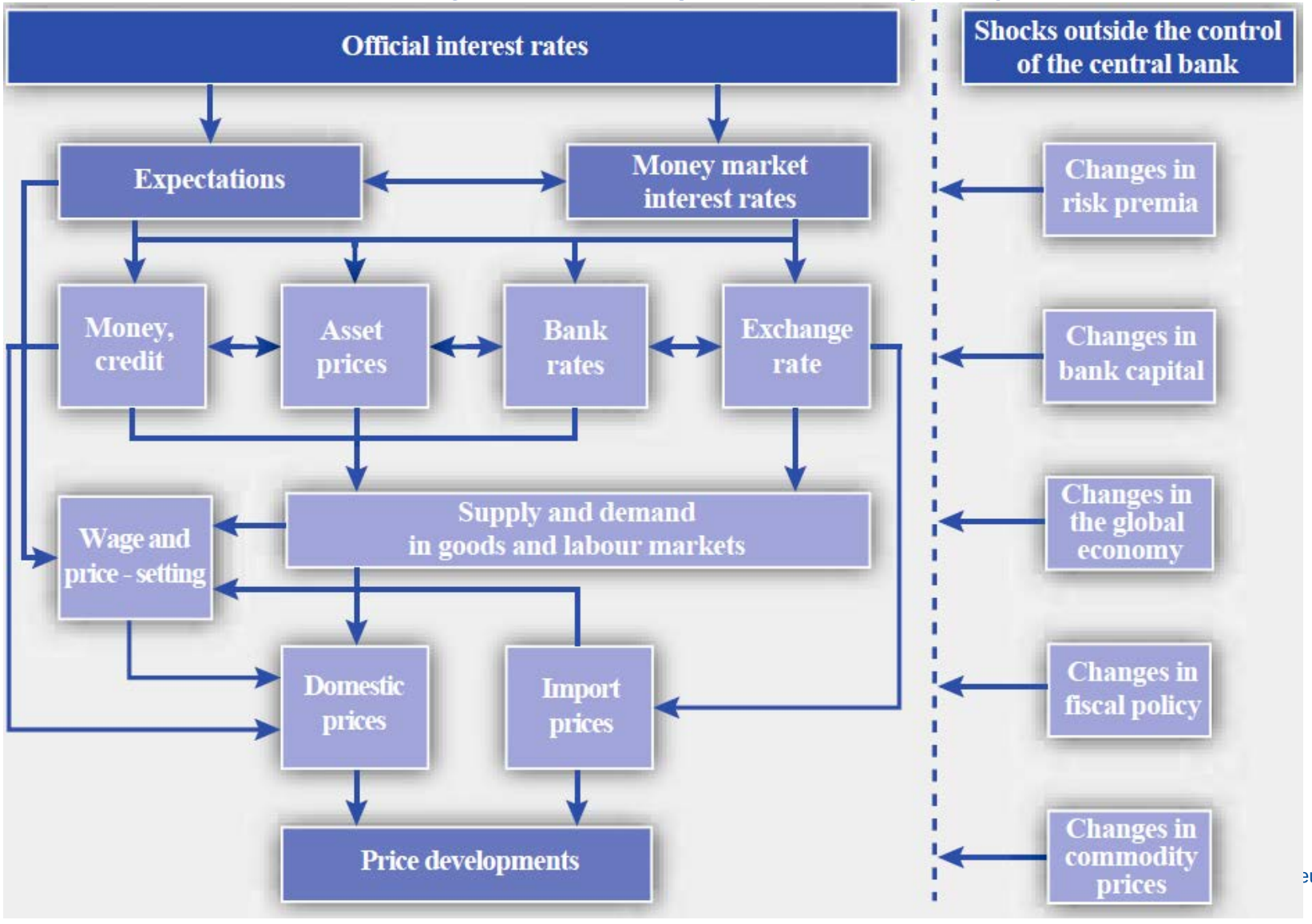
Q&A

Annex: Examples of Fintech outreach in finance and Fintech companies
European Central Bank Paper Series - An outlet for big data statistics and research

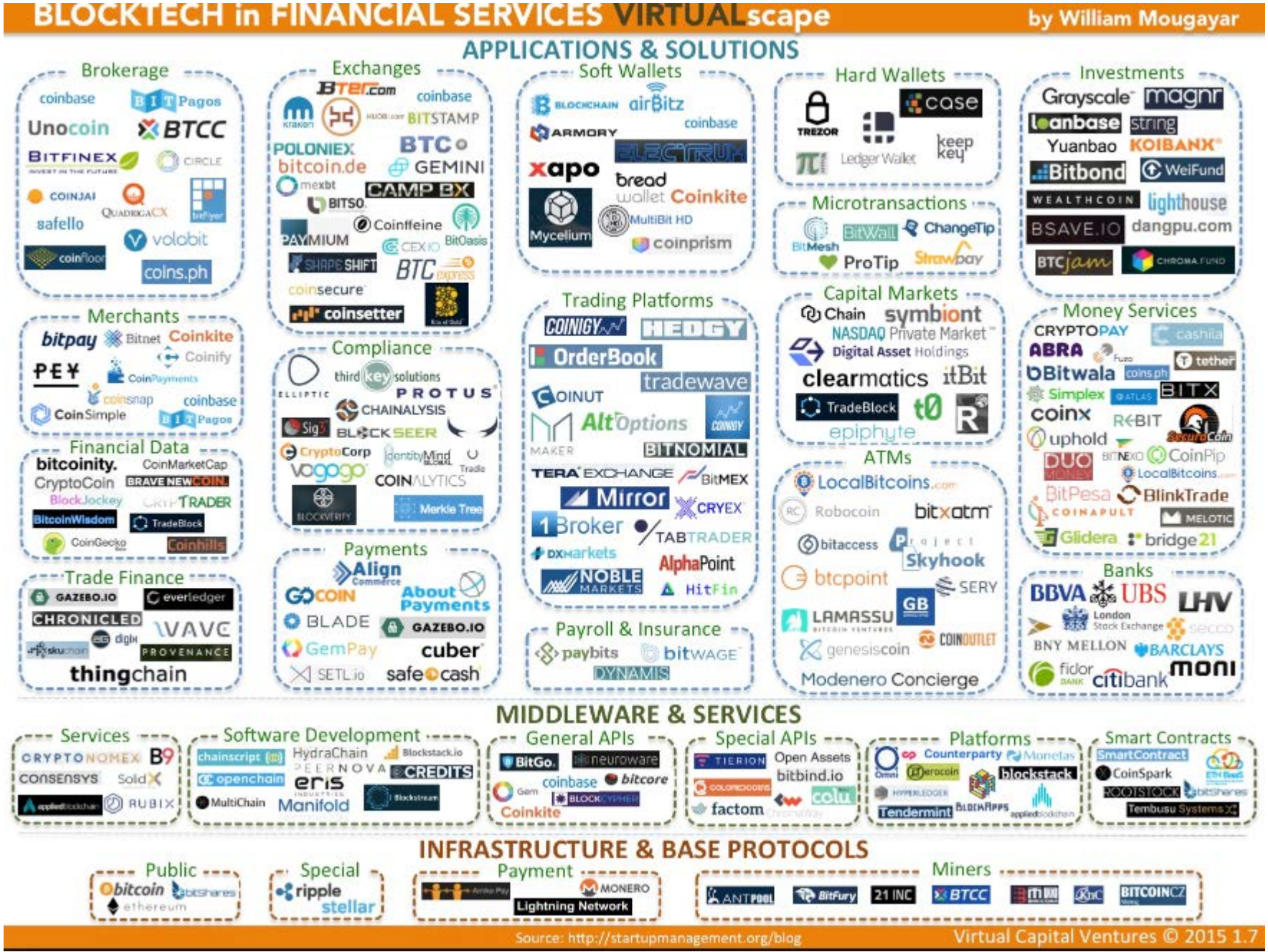
The monetary policy transmission process

How does monetary policy influence price levels?

"Transmission process", i.e. the process through which actions of the central bank are transmitted through the economy and, ultimately, to prices

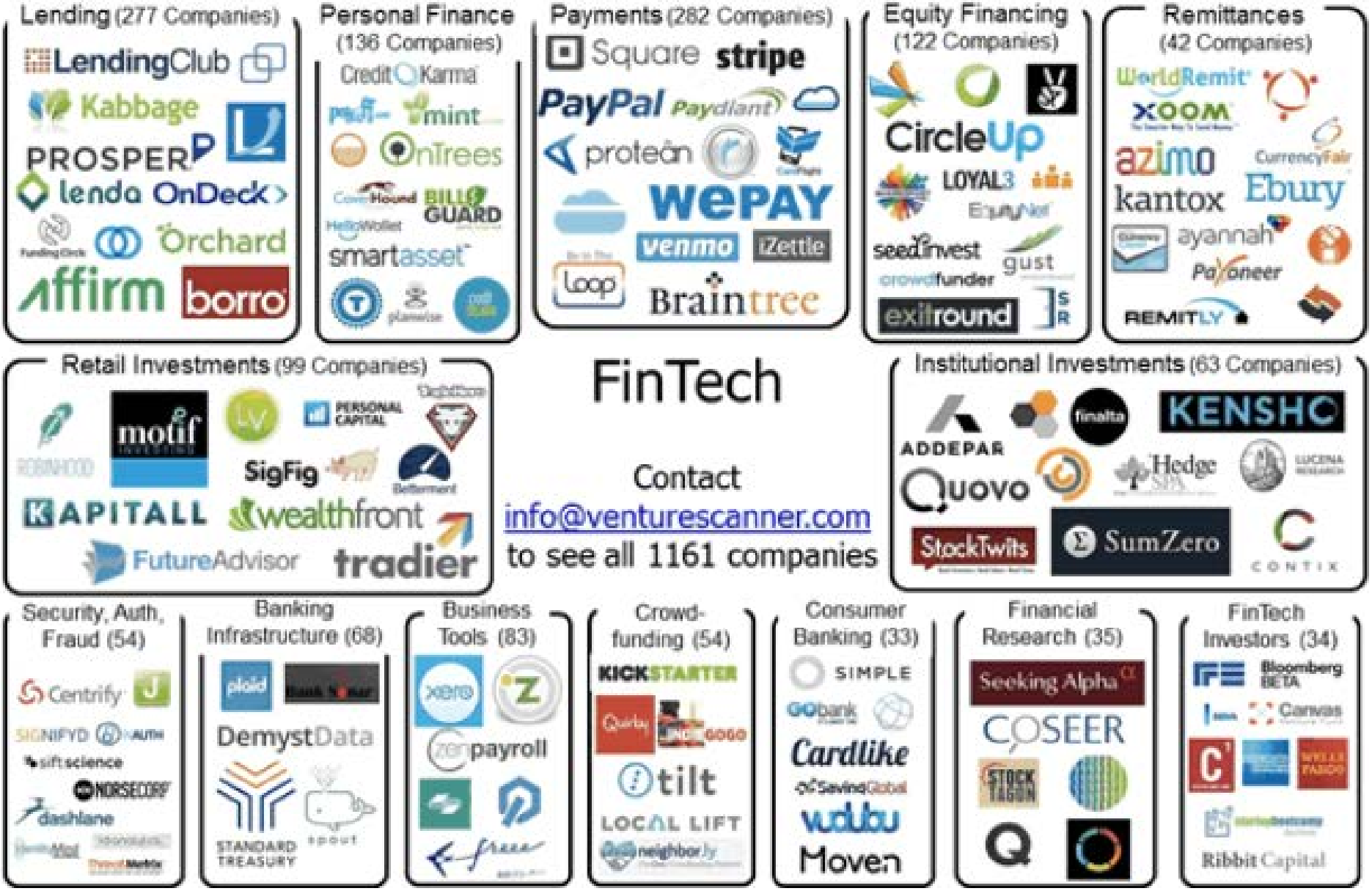


Annex : FinTech in financial IT services – growing field



Annex : FinTech outreach in finance – growing field

Annex : List of Fintech companies – growing field



Annex: European Central Banks Paper Series

➢ An outlet for big data statistics and research:

- **"Nowcasting GDP with electronic payments data" by Galbraith J & Tkacz G.**
  - Electronic payment transactions can be used in nowcasting current gross domestic product growth
  - finds that debit card transactions contribute most to forecast accuracy
- **"Social media sentiment and consumer confidence" by Daas P & Puts M**
  - Relationships between the changes in consumer confidence and Dutch public social media?
  - Could be used as an indicator for changes in consumer confidence and as an early indicator
- **"Quantifying the effects of online bullishness on international financial markets" by Mao H & Counts S, Bollen J.**
  - Develops a measure of investor sentiment based on Twitter and Google search queries
  - Twitter and Google bullishness are positively correlated to investor sentiment