Irving Fisher Committee on
Central Bank Statistics

BANK FOR INTERNATIONAL SETTLEMENTS

IFC – Bank Indonesia International Workshop and Seminar on *"Big Data for Central Bank Policies / Building Pathways for Policy Making with Big Data"*

Bali, Indonesia, 23-26 July 2018

Big data: new insights for economic policy – The Bank of England experience[1]

Paul Robinson,
Bank of England

[1] This presentation was prepared for the meeting. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the meeting.

BANK OF ENGLAND

Big data: new insights for economic policy – The Bank of England experience

Building Pathways for Policy Making with Big Data

BI – IFC International Seminar on Big Data

Paul Robinson, Bank of England

26 July 2018

Outline

- Opportunities
- Challenges

Opportunities

Policy making is an inexact science

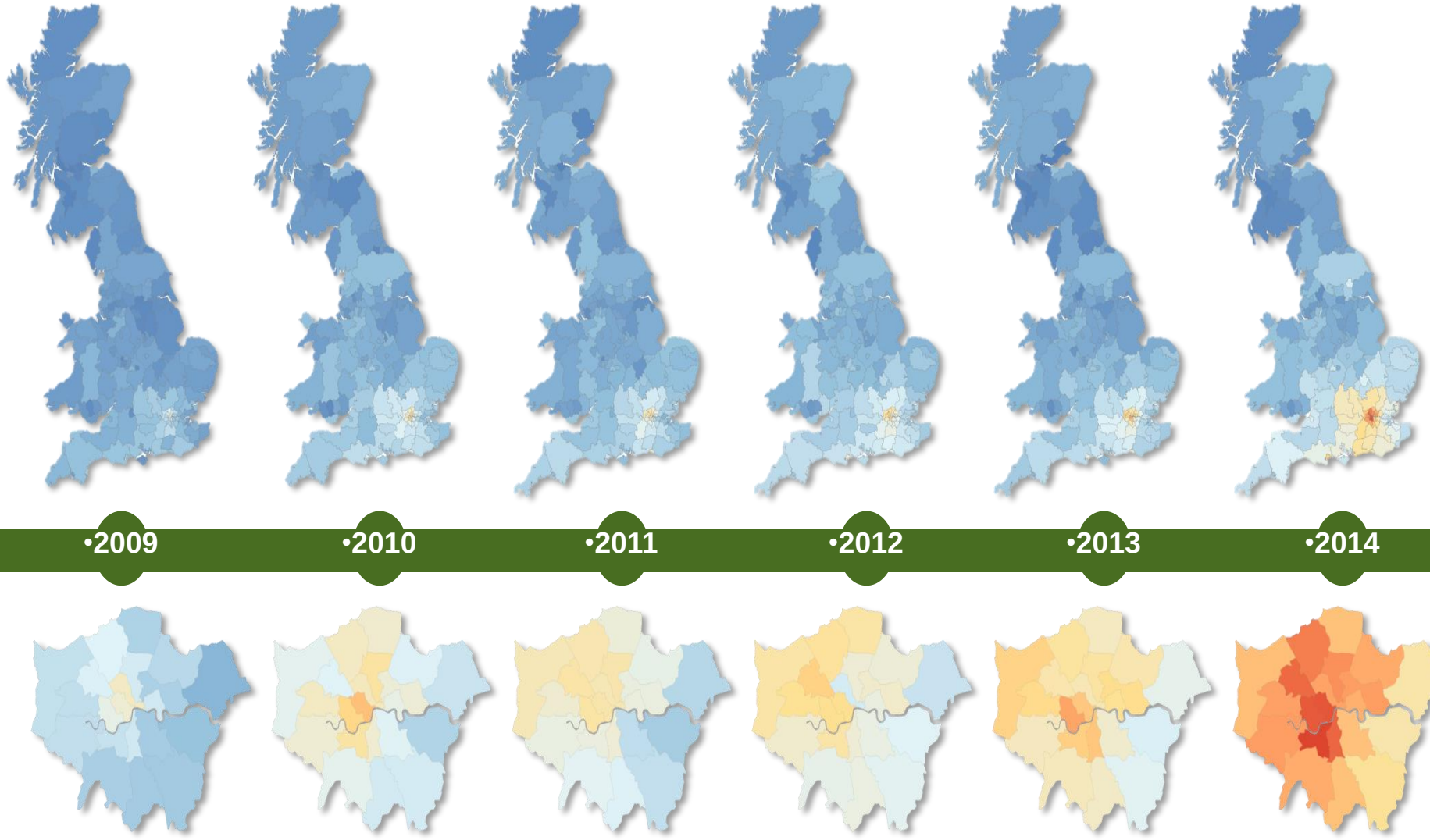
- Imperfect measurement
  - Noise, biases, blind spots, out of date information, (near) simultaneity of cause and effect
- “Too much” data, too little information
- Imperfect theory
- Complex, adaptive system with lots of feedback
  - Leads to “chaotic” behaviour
- Internal frictions

How can Big Data help?

- Imperfect measurement
  - Insight into previously hidden phenomena
  - Combining different types of data
  - Speed and completeness of coverage
- "Too much" data, too little information. Use data science methods to:
  - Improve processing large data sets
  - Help separate the signal from the noise
- Imperfect theory
  - Hypothesis generation
  - Alternative modelling approaches (eg Agent-based models)
- Complex, adaptive system with lots of feedback
  - Difficult to cope with, but more accurate understanding of initial conditions and more frequent updating help a lot
- Internal frictions
  - Improved management information

Big data sets offer significant potential advantages

- Greater **detail** (Volume, Velocity, Variety)
- Allow insights that aggregate numbers might obscure
- Examples:
  - UK housing market
  - Market dynamics around the abolition of the EUR/CHF floor
  - Market liquidity around large market moves

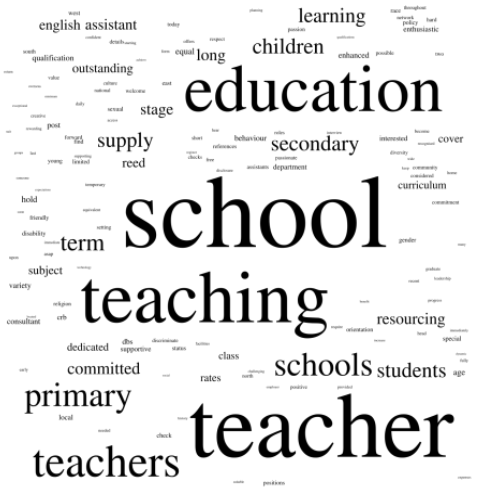


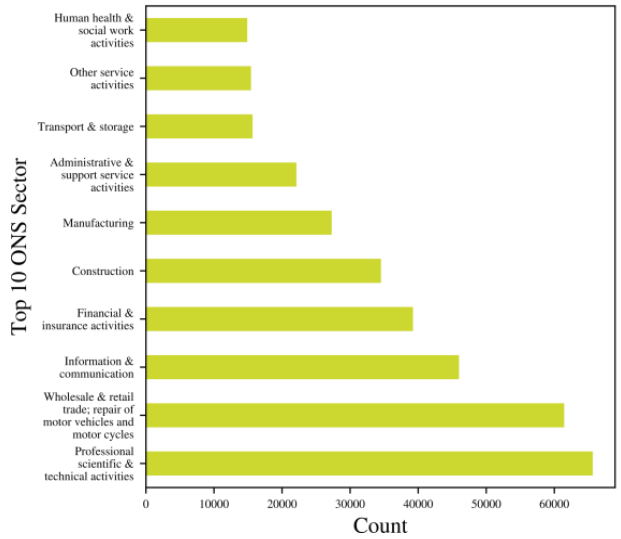
•**Key:** % of mortgages with loan more than 4.5x income

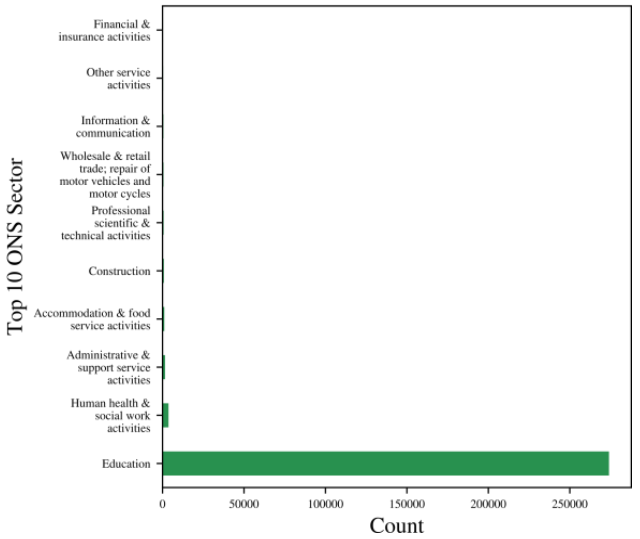
Big data sets offer significant potential advantages

- Greater **flexibility** (Velocity, Variety)
- Gives a window into changing structure of the economy
- Example:
  - Using job adverts to understand changing labour market dynamics

Understanding the labour market using job ads







Big data sets offer significant potential advantages

- Greater **timeliness** (Velocity)
  - ‘Nowcasting’ and ‘nearcasting’
  - Always important, especially in times of crisis
- Greater **efficiency** / value for money (Value)
  - Using administrative data
  - ‘Found’ data

Using Machine learning

- Best used as a complement, rather than a substitute, for other methods
- Objective:
  - The data speak more directly
  - Avoids “unbelievable assumptions of convenience” that underly some alternative modelling strategies
- Inductive rather than deductive approach
- Particularly well-suited to prediction problems
  - Eg statistical issues
  - Nowcasting

The topics

Word clouds of topics found using Latent Dirichlet Allocation.

Topic 0

Topic 1

Topic 2

Topic 3

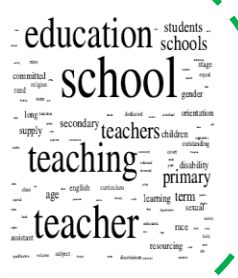
Topic 8

Topic 9

Topic 10

Topic 11

Topic 4



Topic 5

Topic 6

Topic 7

Topic 12

Topic 13

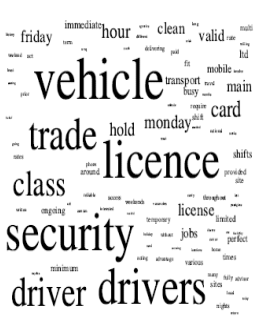
Topic 14

Topic 15

Topic 16

Topic 17

Topic 18



Topic 19

Challenges

Lots of data is not the same as lots of information

- Central banks are typically particularly focussed on avoiding:
  - Monetary dysfunction:
    - High inflation
    - Turbulent currency movements
    - Liquidity traps
  - Financial instability
  - Bank failures
- None of these happens frequently
  - So we often lack a track record

Lots of data is not the same as lots of information

- ML models are often black boxes
  - So policy makers find it hard to ‘inspect the mechnism’
- Together these make it difficult for policy makers to have confidence in the stability of ML models and predictions
- Lots of data does facilitate lots of (potentially spurious) correlations
- And lots of models
  - Similar forecasting ability but very different implications
  - How do we make the right choices?

Correlation versus Causality

- ML focuses on prediction
  - Not on structural models
  - But central banks set policy and a policy intervention may change the structure of the economy
  - Beware the ‘Lucas critique’ (and structural breaks)
- This does not mean that ML is not a good fit for central banks
  - Forecasts often matter
  - Intermediate targets can be useful

“Veracity”

- Big data sets are often populations, not samples
  - Therefore no sampling error
- But the observed population characteristics may not be typical of the underlying data generating process
- Or it may be biased relative to the true population of interest

Confidentiality / ‘Big Brother’ state

- This was not relevant to the CPI work
- In general, the more detailed and granular the data set is, the more likely it is to contain confidential information
- We must ensure that:
  - we only use data for appropriate reasons
  - the minimum number of people are able to see any confidential data given the needs of the situation
  - data are stored securely and professionally

Conclusion

- Big data has much to offer
- But it is not, and likely never will be, a panacea
- Policy judgement will always remain crucial
- So we need to ensure that we move towards:
  - Combination of the best parts of structural and data-driven models
  - Interpretable outputs from big data models