Data-driven economic models: challenges and opportunities of big data

Monica Bulger, Ralph Schroeder,[1] and Greg Taylor [1]

[1]Oxford Internet Institute, University of Oxford

NEMODE
26 November 2013

Monica Bulger

Ralph Schroeder

Greg Taylor

Introduction

Some sectors and firms characterised by rapid incorporation of big data into business models and practices.

Many parts of the economy have been slow to adopt, despite evidence of advantages.

Introduction

Some sectors and firms characterised by rapid incorporation of big data into business models and practices.

Many parts of the economy have been slow to adopt, despite evidence of advantages.

Two main questions:

- What are the main obstacles to the broader use of big data in the UK economy?
- What can we learn from others (sectors/countries/firms) about best practice in removing these barriers? Is there anything policymakers can do to help?

Methods

- Qualitative research building on interviews with people from the business/policy world.
- Sectors: Advertising, public services, data management & analysis, risk & finance, GIS, human resources, online services, retail, software, research/education, biotech.
- Example organisations: IBM (IT solutions), Gild (recruiting), Willis Re (insurance), Acxiom (data), Proteous (biotech), Datacratic (real time consumer data), Brightscope (financial information).

What is big data?

Not a lot of consensus

- Most see this as an evolution rather than a revolution.
- Scale of data not very interesting because systems can often scale easily.
- Challenges (and opportunities) around new *kinds* of data (e.g. non quantitative) and also ways of handling real-time data.

What is big data used for?

We see a broad range of uses for big data, including new business models:

- Data as a service
- 'Solutions'
- Analytics outsourcing
- Consultancy

What is big data used for?

We see a broad range of uses for big data, including new business models:

- Data as a service
- 'Solutions'
- Analytics outsourcing
- Consultancy

and revision of existing practices

- More sophisticated business intelligence
- New products (e.g. targeted advertising, smart vehicles)
- Refining business processes (e.g. better risk assessment, better tracking of ROI).

What is big data used for?

We see a broad range of uses for big data, including new business models:

- Data as a service
- 'Solutions'
- Analytics outsourcing
- Consultancy

and revision of existing practices

- More sophisticated business intelligence
- New products (e.g. targeted advertising, smart vehicles)
- Refining business processes (e.g. better risk assessment, better tracking of ROI).

There is an entrepreneurial opportunity. Little capital needed for some business models.

Data and culture

We get a very clear message that data needs building into core of a business.

- Our subjects say that successful companies make data part of their business plan, rather than an afterthought.
- Design data collection with specific business objective in mind to ensure the right systems and skills are in place.

Data and culture

We get a very clear message that data needs building into core of a business.

- Our subjects say that successful companies make data part of their business plan, rather than an afterthought.
- Design data collection with specific business objective in mind to ensure the right systems and skills are in place.

Also build into organisational structure.

- Don't silo big data analytics: build into existing business intelligence unit and make sure it understands the business imperative/objective.
- Important to have people in senior positions who 'get' data and can work to build it into the business model.

Skills

Not much sentiment that skills are lacking, but hiring is key.

- Not a lot of emphasis on undergraduate level training: Most of the *specific and new* opportunities seem to be for higher level experts who can deal with genuinely new statistical problems.
  - True of big firms, but also the small analytics firms we spoke to were founded by people with postgraduate-level (or above) stats skills.

Skills

Not much sentiment that skills are lacking, but hiring is key.

- Not a lot of emphasis on undergraduate level training: Most of the *specific and new* opportunities seem to be for higher level experts who can deal with genuinely new statistical problems.
  - True of big firms, but also the small analytics firms we spoke to were founded by people with postgraduate-level (or above) stats skills.
  - Some see the novelty as fundamental to what big data is (others question the notion of big data altogether).

Skills

Not much sentiment that skills are lacking, but hiring is key.

- Not a lot of emphasis on undergraduate level training: Most of the *specific and new* opportunities seem to be for higher level experts who can deal with genuinely new statistical problems.
  - True of big firms, but also the small analytics firms we spoke to were founded by people with postgraduate-level (or above) stats skills.
  - Some see the novelty as fundamental to what big data is (others question the notion of big data altogether).
- Emphasis on building a team that brings together different skills to solve specific problems.
  - Need a range of competencies to talk to each other to make big data work for a well-defined business objective.

Scientific and academic dimension

General agreement that successful big data companies cultivate an academic/scientific culture among their analysts.

- Big data can prove anything, so need to be honest.
- Respect for scientific method: design experiments that can fail and be willing to be proven wrong.
- Given rapid pace of change, need intellectual freedom to explore and adapt (and workers who are capable of adapting).

Public/private data

Wide variety of different types of datasets, including public and open, proprietary, data scraped from web.

Public/private data

Wide variety of different types of datasets, including public and open, proprietary, data scraped from web.

Low hanging fruit is quantitative data, but also big opportunities for more exotic data through novel analytic techniques such as machine learning or sentiment analysis.

Public/private data

Wide variety of different types of datasets, including public and open, proprietary, data scraped from web.

Low hanging fruit is quantitative data, but also big opportunities for more exotic data through novel analytic techniques such as machine learning or sentiment analysis.

Publicly available or open data rarely used in isolation, usually combined with semi-public, or private, proprietary datasets.

- e.g. Willis Re use of Lansat images and weather data combined with customer information for risk analysis.
- e.g Brightscope use of publicly available finance records combined with proprietary demographic data to recommend retirement investments.
- This is a key way in which value is added to public data (the other being curation).

Standardisation and compatibility

Many of the advantages of big data stem from combining data or using it in novel ways.

- A big challenge for firms is working with data and systems that are not standardised or compatible.
- Introduces inefficiency as interoperability has to be engineered in.

Standardisation and compatibility

Many of the advantages of big data stem from combining data or using it in novel ways.

- A big challenge for firms is working with data and systems that are not standardised or compatible.
- Introduces inefficiency as interoperability has to be engineered in.
- Also a problem of short-termism: lazy use of metadata or lack of thought for future compatibility renders legacy data/systems useless.
- Links to earlier point about having specific objectives in mind for data and making that core to the organisation.

Standardisation and compatibility

Many of the advantages of big data stem from combining data or using it in novel ways.

- A big challenge for firms is working with data and systems that are not standardised or compatible.
- Introduces inefficiency as interoperability has to be engineered in.
- Also a problem of short-termism: lazy use of metadata or lack of thought for future compatibility renders legacy data/systems useless.
- Links to earlier point about having specific objectives in mind for data and making that core to the organisation.

Many respondents identified lack of standardisation of public data as a key obstacle to its use.

- Often in non machine-readable form.
- One respondent says 80% of their time is cleaning data.

Privacy a key issue

Many respondents identified privacy policy as a key stumbling block.

- A sense that there is not a level playing field.
  - Some sectors, companies seem to be able to get away with more.
  - There seems to be a taboo on third party data use.
  - Linked to consumer understanding of and attitude toward data use.
- Germany has been praised for privacy laws that are transparent, consistent, and adapt to evolving data practices.
- Also a need for harmonisation of policy across borders.
- On a related note: many firms ill-prepared for the ethics of handling personal data and ensuring its quality.

Public incentives

- Some countries have in place public incentives to grow the data analytics sector.
- Few respondents mentioned this as a valuable way for the government to intervene.
- Analytics should be seen as a means to an end, rather than an end in its own right.