Utilising Big Data to Improve Economic Statistics

Opportunities and risks using Big Data and machine learning

Centre for Central Banking Studies, Bank of England, 26-27 November 2018

Rebecca Riley, ESCoE and NIESR

ECONOMIC STATISTICS CENTRE OF EXCELLENCE

A COLLABORATION WITH Office for National Statistics

Opportunities and Risks

Timeliness

Granularity

Taxonomies

Survey weights

New data sources & statistics

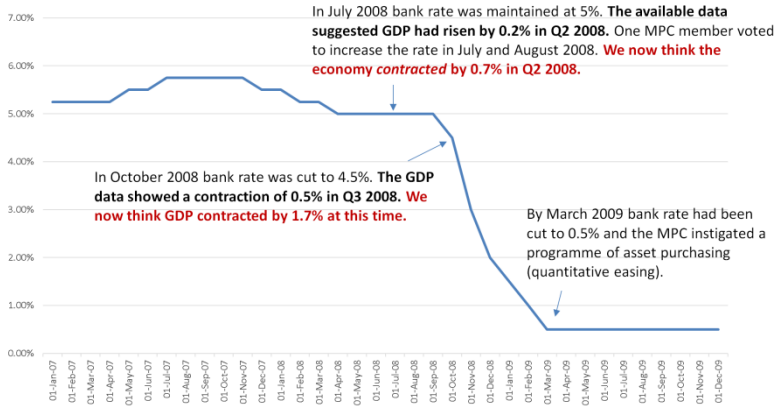
Representativeness

Data continuity

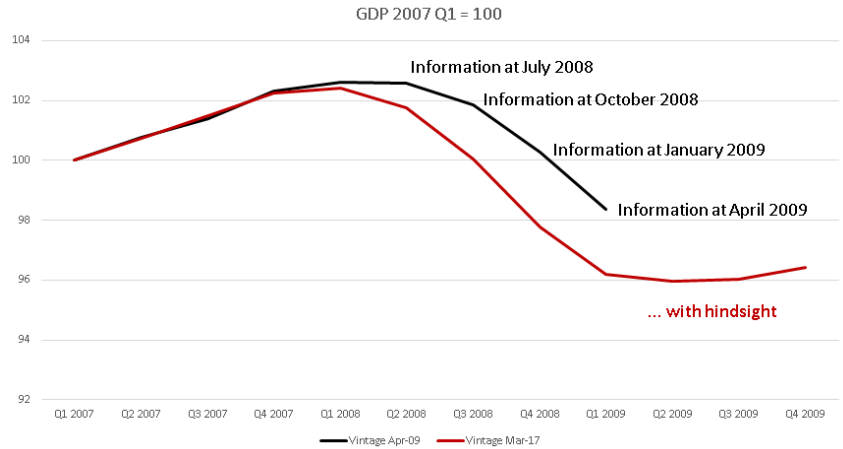
Increasing the timeliness and accuracy of GDP …

New opportunities using VAT data and other Big Data

Bank Rate during the Great Recession



GDP during the Great Recession



… better decisions

ECONOMIC STATISTICS CENTRE OF EXCELLENCE

A COLLABORATION WITH Office for National Statistics

Timely and high frequency data for the UK regions

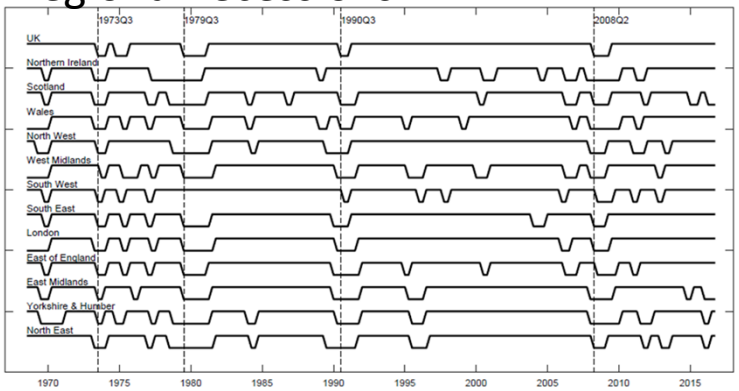
Annual & quarterly data used to "fill in the holes" in the current regional database, using methods adapted from the machine learning literature.

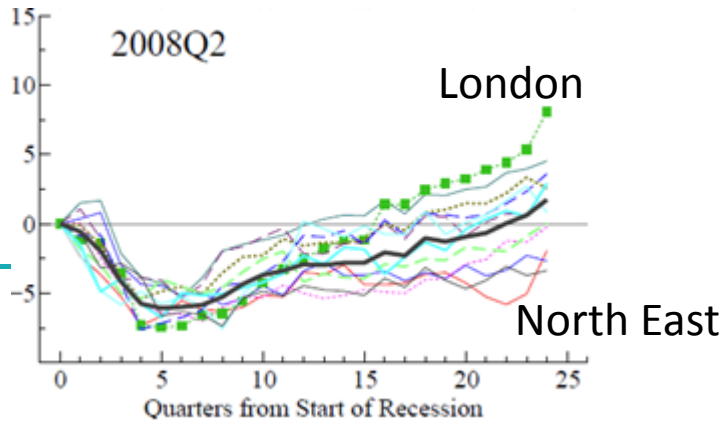
Allow for interactions and consistency with aggregate statistics.

Possibilities: lower level geographies

Regional Output Growth in the UK: More Timely and Higher Frequency Estimates, ESCoE DP 2018-14 (Koop, McIntyre, Mitchell and Poon, 2018)

Regional recessions





ECONOMIC STATISTICS CENTRE OF EXCELLENCE

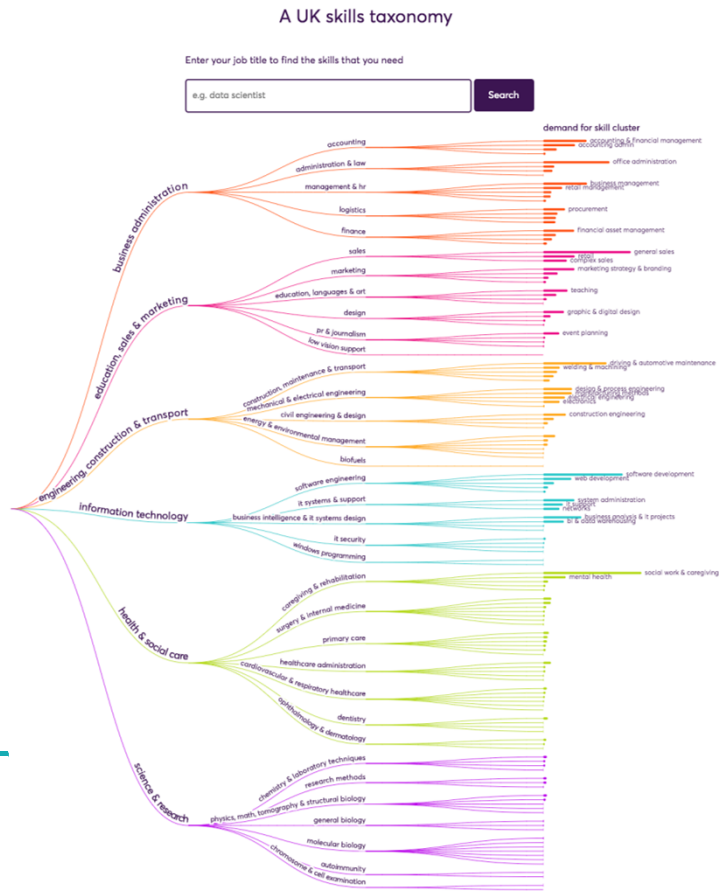
Improving understanding of the labour market

Apply machine learning techniques to job adverts to measure skill co-occurrence and similarity, complementing the SOC

Limitations: incomplete, hierarchical

Possibilities: can study emergence of new skill demands

An Open and Data-driven Taxonomy of Skills Extracted from Online Job Adverts, ESCoE DP 2018-13 (Djumalieva and Sleeman, 2018)



ECONOMIC STATISTICS CENTRE OF EXCELLENCE

www.escoe.ac.uk