THE IMPACTS OF OPEN DATA: TOWARDS EX POST ASSESSMENT

Heli Koski
The Research Institute of the Finnish Economy
25/11/2015
Share-PSI 2.0 Workshop, Berlin

"The open-data revolution has not lived up to expectations. But it is only getting started"
(Economist, Nov 21st 2015)
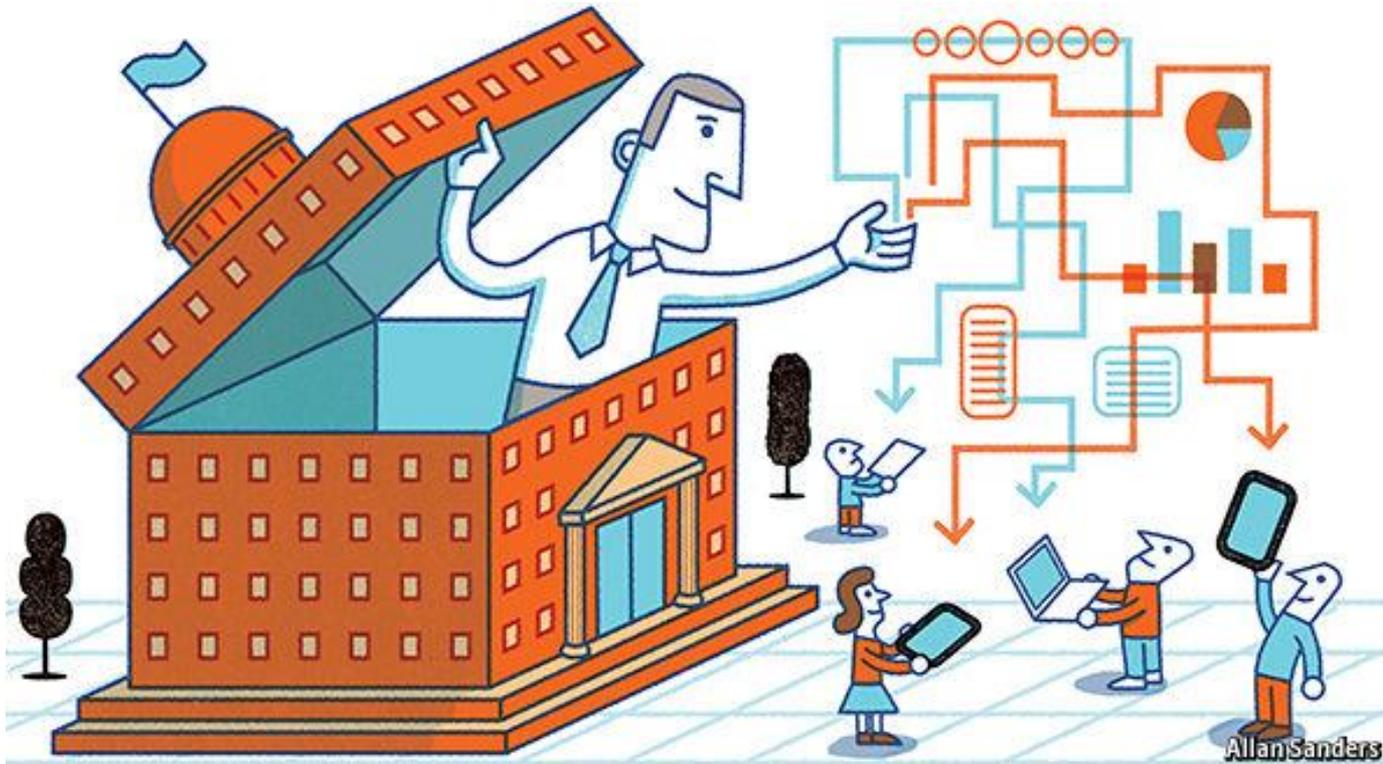
"The open-data revolution has not lived up to expectations. But it is only getting started"
(Economist, Nov 21st 2015)

- The article gives bunch of open data examples (health applications, transport timetables, crime logs, pollution readings etc.) improving peoples lives but asks: "*why more has not been achieved*?"

- Provides four answers to that

i) Data made available often useless.

ii) Data engineers & entrepreneurs who might be able to turn it into useful, profitable products find it hard to navigate.

iii) Too few people capable of mining data for insights or putting the results to good use.

iv) It has been hard to overcome anxieties about privacy.

The impacts of open data – where we stand?

- The Open Data Barometer project reported that, in 2013, none of 77 countries studied had conducted comprehensive assessment of the impacts of opening up data.

- Various examples suggest that opening up government data significantly increases data use

  - does not reveal anything about the socio-economics impacts of open data use.

The impacts of open data – where we stand?

- Socio-economic research of open data use relatively scarce; very little known about the economic implications of data use at the level of organizations or the economy as a whole.

Opening up data – what we know?

i) *Impacts on firms and industries*

- Innovation, growth, efficiency, productivity?
- Various studies suggest that opening up of public data may facilitate innovation and growth, particularly in SMEs
- Size and growth of certain data-based markets closely linked to the openness of data used

ii) Impacts on citizens

- Time savings (e.g. when rush hours are avoided due data use), monetary savings from free access to data, better decisions and participatory democracy.

Opening up data – what we know?

iii) *Impacts on the public sector*

- Improving the efficiency and quality of services and cutting costs
- Increased tax revenues

iv) *Economy-level impacts*

- Employment, growth, productivity?
- *Ex ante* estimates: economic potential of open data remarkable but reported estimates vary greatly in size.
- Only few innovations (electricity, steam, ICT) have had significant macroeconomic impacts, which have typically been realized decades after the initial adoption of innovation.

Towards ex post evaluation of the socio-economic impacts of open data

- Ex post analysis focusing on materialized economic impacts is still in its early stages.
- Challenging research topic due to measurement difficulties, lack of systematically collected statistical data on data use & phenomenon still relatively new.

- *Research* – will require systematic, long-term efforts.

- Understanding impacts of data use important to ensure that maximum socio-economic benefits will be gained from open data use.

- Prerequisite for research concerning the impacts of open data is *development of an assessment model of impacts.*

Development of monitoring and assessment model

- The key questions:
  i)   What are potential indicators for impact assessment (measuring expected key impacts of open data)?
  ii) What useful data is already being collected?
  iii) What additional data are required and how to collect it?
- Basis for detailed development of monitoring model and set of indicators.
- Mainly covers measurable, economic impacts, but can be complemented by qualitative assessment.

Key economic impacts
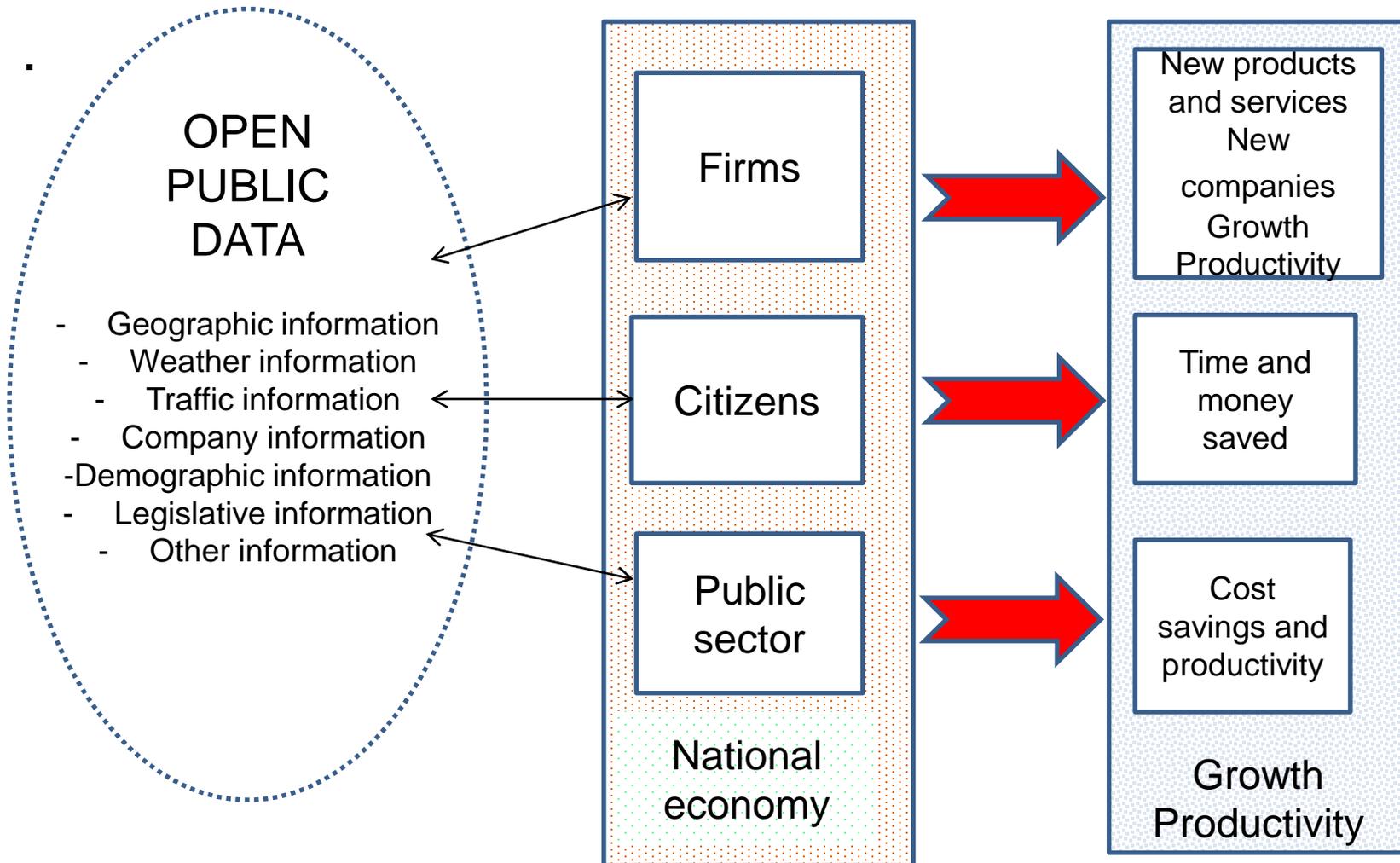
Indicators for expected economic impacts

	Expected impact	Indicators	Potential sources of information
Companies	New products and services New companies	The quantity of new products and services based on open data The share of income from open data services/products of total turnover	Surveys, adding questions to e.g. the "Use of information technology in enterprises" survey & and Eurostat's Community Innovation Survey.
	Growth	Changes in turnover and number of employees	Business registers of national statistical offices Firm-level financial statements databases
	Productivity	Added value per employee or per hour of labour	Firm-level financial statements databases.
Citizens	Time and money saved	Amount of free data used x earlier price Estimated time savings	Surveys, adding questions to e.g. the "Use of information and communications technology" survey.
Public sector	Cost savings Productivity	Changes in labour hours x labour costs per hour Service output in relation to the resources used	Automatic monitoring of information and service offering Surveys, the Finnish Government ICT Review, adding questions therein Statistics on central government productivity
National economy	Growth Productivity	Changes in added value i.e. GDP GDP per capita	Databases of national statistical offices.

What useful data currently collected?

- Indicators for firm-level impacts available from financial statements databases and databases of national statistical offices
- Economic impact indicators at the level of economy also from national statistical offices

- Currently no data on the role/use of open data in firms, household and public sector is systematically collected.
- Primary means to get data surveys, e.g., by adding open data questions to mandatory, regular surveys among EU-countries (+ national surveys)

Data collection needs: citizens

- Indicators:

- Impacts: time and money savings due to open data too difficult to assess

- Data use: types of open data used and to which purposed data are used.

- Questions, e.g., to the annual EU-wide "Use of information and communications technology" – survey.

Data collection needs: firms

- Indicators on data use:
  - Extent and purpose (e.g. innovation, planning of logistics) for which firms use open data and types of data used.
  - Role of open data in the establishment of new companies, their operations and job creation

- Questions to the annual "Use of information technology in enterprises" –survey and biannual CIS (Community Innovation Survey) –survey.
- Both surveys EU-wide

Data collection needs:
public sector & economy-wide

- Indicators for public sector

- Impacts: Costs (amount of additional work needed or time saved by opening up data) & productivity

- Data use: Amount of data shared by public administration organizations; extent of data use and data flows between public organizations

- Questions to the national surveys?

- At the level of economy

- Supply and use of open data can be measured, e.g., by quantity of opened-up data, number of users and volumes of use.

Ex-post impact assessment

- After data collection, (partial) *cost-benefit analysis* can be conducted to assess welfare implications of opening up of public data.

- Aim is to define monetary value for all impacts of open data and use this to calculate their total value and profitability for society.
    - In practice, key impacts ("minimum threshold" of impacts)
- In addition: qualitative impact assessment

Ex-post impact assessment

- Survey data on use of open data can be combined with various statistics such as firm-level financial statement databased and business registers of national statistics offices.

- Long-term monitoring and compilation of statistics on open data use would enable quantitative analysis, using advanced statistical methods, of the economic impacts of open data.

- Surveys could also be used to assess other social impacts of open data. Qualitative research methods best suit for this.

Other measures and actions

- In addition: user experiences of citizens, firms and public sector organizations in order to assess usefulness of different public data resources.

- Natural platform for user-driven assessment and feedback are open data platforms.

- Also: development of means to generate information on best practices concerning open data use & efficiently disseminating best practices for use within different organizations.

Conclusions

- Before asking "*why more has not been achieved*?" we would need to understand "*how much has been achieved*"

- To answer this question we need systematic, long-term efforts to be able to evaluate impacts of open data use:

i)   Development of assessment model with proper indicators

ii)   Systematic collection of data

iii)   Research.