

Maximising Returns to Government Investment in Data: Data Management

Tanvi Desai
London School of Economics

Value of Microdata

- Government focus on the aggregate /macro results

HOWEVER

- Data once lost can never be recaptured
- Can investment in data collection be justified if only used to create limited aggregate figures?

Records

- All government departments should keep a comprehensive record of data collected.
AT THE VARIABLE LEVEL!
- Ideally - an independent governmental body with responsibility for holding and sharing information about government data resources (would not necessarily need to hold data)

Metadata

- A consistent metadata scheme agreed and implemented by all government departments
- Must include information on context in which data collected

Archiving

- If the data has value for answering questions now - it will be valuable for answering questions in the future
- Confidentiality not an excuse as not relevant forever cf. Census release after 100 years

THEREFORE

To maximise investment all data must be archived.

Contractors

- Any person/organisation contracted to collect data for government should deliver the microdata - not just a report
- Common argument that passing data to government reduces response rates
- Even if true (?) investment in data collection must always be weighed against utility of data now AND IN THE FUTURE

Sharing

- Data sharing between gov. depts. should be facilitated
- Reduce collection of duplicate variables
- Minimise respondent burden
- **SAVE MONEY!**

Linking

- International interest in effective methods of data linking. UK specific - Beyond2011

Confidentiality concerns

- Need for trusted disinterested party(ies) to carry out data-linking
- Is a national index for linking possible? ID cards stealth?

Consent to link

- Is it necessary to secure consent to link from respondents?

If yes then consent to link should be included for all data collected BUT

- Is it necessary to secure consent for each type of link eg. medical, economic etc.
- How does it impact response rates/sample bias

How do we pay for it?

- Reduced cost of data collection due to reduced chance of duplication
- Reduced cost of data collection due to data linking
- Reduced cost due to centralisations no need for duplicate infrastructures

How do we pay for it?

USE EXISTING SERVICES

NO NEED TO REINVENT THE WHEEL

Many countries have excellent existing infrastructure in academia and government invest in those and avoid high start up costs

THANKYOU

Tanvi Desai
t.desai@lse.ac.uk