Access to privately held data – Australia’s experience
Recap on working paper 1

Expert Meeting on Modernizing Statistical Legislation
18 November 2021

George Holton, Program Manager
Data Strategy, Integration & Services
Australian Bureau of Statistics

Australian Bureau of Statistics
Informing Australia’s important decisions
Our experience before COVID-19

- Supermarket scanner data (transactions)
- Introduced 2012
- Large supermarkets (Australia has 4)
- Represent 25% of weighted prices which make up our consumer price index
- Provided voluntarily, under ongoing written agreements
- Preferable to supermarkets than previous price data collection arrangements
- Detailed information about transactions generated from interactions between customers and businesses – generally more detailed than survey data from businesses
New statistical requirements

- Economic uncertainty
- Need for up to date information on economic responses of individuals and businesses
- Need to understand patterns and trends in personal movement in real time
- Need it fast

Attention quickly turned to micro data held by the private sector

- Largest financial institutions (Australia has 4 main banks)
- Largest retail businesses
- Electricity and communications data
The Australian Statistician contacted CEOs to seek support

Strong and quick support from major companies

Two major topics to be resolved:

(1) data supply arrangements
(2) privacy and confidentiality of data
Privacy and confidentiality

We only acquired de-identified, aggregated microdata from the private sector

Dealt with under existing law

Confidentiality of the data protected by our legislation

We committed to each company in writing to:
(1) only use their data as input to aggregated data
(2) never release to a third party
Data supply arrangements

- Voluntary - not compelled by law
- Mix of existing pro forma data supply agreements and licence agreements
- We had to seek legal advice on intellectual property and financial liability
- We agreed to accept data in whatever format it was in
- We agreed to use whatever secure technology the business used for data transfer
- Agreements in place within 2 months of pandemic hitting Australia
Uses and benefits

- National statistics on inflation, business sales, household consumption and gross domestic product
- Near real time consumer and business responses to the pandemic – supply chain resilience
- More timely than existing statistical publications
- More granular levels of data
- No provider burden for individuals, businesses or households
- Close to real time insights into personal and household movement
- Used by governments to shape support policies
Normalising this data

- Upfront flexibility on data formats and transfer technology = fast receipt during the pandemic

- But leaves operational challenges to be normalised now:
  - How we partner with the private sector on an enduring basis
  - How we adapt our organisational capabilities to manage and use privately held micro data
  - How we adapt our business to a more balanced used of private micro data beyond survey supplementation
To maintain voluntary settings, health of our relationship with private providers is key. How do we do that well?

What level of private sector data dependency and supply security risk are we ok with?

Private companies are offering future data supply and data access models if we want their data. Is that ok?

How can we reduce the provider burden on private providers across all of the data types we ask them for?

How can we streamline and standardise the data supply agreement and governance processes?

How can we standardise the data supply pipeline across all providers?
Adapting our capabilities

- Use of cloud based solutions to ingest, store and process data at scale
- Building the capabilities to deal with higher frequency data
- Using the data as an organisational wide resource – not an input to just one survey or product
- Standardising many bespoke supply arrangements and technology solutions into standard arrangements used by all
- Clarifying internal roles and responsibilities for data supply negotiations and provider relationship management
- Adequately funding data management and technology needs
Thank you