Making Greater Use of Administrative Data in Economic Statistics

Jeff Cope
Statistics New Zealand

May 2016
Potential benefits from increased use of administrative data

• Meeting customer needs
  – Statistical output and production flexibility
  – Scalability in statistical production
  – Improved quality
  – Microdata availability

• Efficiencies
  – Internal production efficiencies
    » reduced data collection & processing costs (post-out down > 50%)
    » data reuse for new outputs
  – Reduced respondent burden
Moving to an ‘administrative data first’ paradigm

Key statistical issue
Where to fit the boundary to maximise the use of admin data while maintaining quality.
From survey based to admin based
the Annual Enterprise Survey

(i) AES 1997
Survey based
Full coverage
Postal
Sample Stratum
IR10 data
Sample size: 15% of the population
IR10 data for 230,000 businesses

(ii) AES 1999
Individuals and Partnerships
Full coverage
Postal
Sample Stratum
Sample Stratum
Sample Stratum
IR10 data
Sample size: 6% of the population
IR10 data for 290,000 businesses

(iii) AES 2009
Income threshold
Full coverage
Postal
Sample Stratum
Sample Stratum
Sample Stratum
IR10 Data
Sample size: 5% of the population
IR10 data for 300,000 businesses

(iv) AES 2010
Reduced high weight
Full coverage
Postal
Sample Stratum
Sample Stratum
Sample Stratum
IR10 Data
Sample size: 4% of the population
IR10 data for 310,000 businesses

(v) 2012
For selected industries
Full coverage
Postal
Sample Stratum
Sample Stratum
Sample Stratum
IR10 Data
Sample size: 3.6% of the population
IR10 data for 310,000 businesses

Goal
Admin. data based
Key units collection
The new design for sub-annual economic statistics

From 2001 to 2015

- Large Units: Survey data (Sample (Survey data))
- Small Units: Administrative data (GST sales)

New design

- Variables: Survey data (Managed collection)
- Administrative Data (GST sales)

New design

- Variables: Survey data (Managed collection)
- Admin. Data (GST sales)

Key statistical issue
How to estimate missing variables: using variable relationships in managed stratum, or from relationships to other admin data variables or a small sample.
Key pillars supporting the greater use of administrative data

• Managing the risks in the supply of admin data
  – Legal support, MOUs, and good relationship management ensure consistent and regular data delivery: increasingly seen as a data partnership.

• Comprehensive business register
  – Relationship links between statistical units and collection (survey and admin.) units underpin the strategy and support integration of data.

• Innovative statistical production systems
  – Single processing platform containing all survey and administrative data linked via the BR “spine”.
  – Flexibility and scalability in production and analysis: macro-meso-micro views.

• Administrative data quality framework
  – Assessment framework at design and monitoring phases.