

15th Meeting of the Ottawa Group

10 – 12 May 2017

Session 6: Challenges of “big data”

Towards a big data CPI for New Zealand

Alan Bentley, Statistics New Zealand
Frances Krsinich, Statistics New Zealand

In our digital age, what's the best way to measure inflation? We find an abundance of new data sources. Some of these will prove useful, given the right mind-set and methods, others may prove distracting or prohibitively hard to obtain access to. In this paper we take stock and explore the path to a 'big data' Consumers Price Index (CPI) for New Zealand.

We discuss the benefits to a big data approach, reflecting on our early adoption of model-based approaches to price measurement, such as using a hedonic model for second-hand cars, and retail transaction – scanner data – for consumer electronics products (including TVs, computers, digital cameras) in the Consumers Price Index. A big focus of our work in this area has been the development and theoretical justification of a regression-based approach called 'the FEWS index' that will produce non-revisable quality-adjusted indexes even when there is no explicit information on product characteristics.

We look towards the future by considering the opportunities that are currently in front of us. Notably, we recently signed up to the purchase of a year's-worth of daily web-scraped online price data from PriceStats, the commercial counterpart of MIT's Billion Prices Project. This data captures, in real-time, online prices for a wide range of different NZ retailers and it will enable us to do detailed research on the potential for enhancing and improving our current data collections and price measurement. Using the online data in combination with expenditure information from surveys or scanner data presents a rich opportunity for more frequent and timely price indicators than are currently available.

Preliminary research on measuring rent price change from administrative data will also be used to illustrate our approach to the art of using 'big data'. This data source highlights the opportunities and obstacles of coverage, timing and quality adjustment.