Making Greater Use of Transactions Data to compile the Australian CPI

ACE 2017

Michael Holt
Prices Branch (ABS)

Han Q Pham and Michael Webster
Methodology Division (ABS)

Contents

- Background
- Multilateral methods
- Extension methods
- Framework for assessing methods
- Empirical results
- Consultation
- Implementation

Background

- ABS in a transformation environment – seeking ways to utilise 'big data' for compilation of economic statistics

- *Enhancing the Australian CPI: a roadmap* (ABS 2015) sets out four research priorities
    - Frequency of weight updates
    - **Transactions/scanner data**
    - Monthly CPI
    - Other enhancements

- Transactions data contains detailed information about transactions, dates, quantities, product descriptions, and values of products sold

Product ID	Store Location	Product Description	Time Period	Revenue ($)	Units sold	Unit value ($/unit)
U0001	Sydney CBD	CARROTS PREPACKED 1KG	Jan-16	5000	2500	2.00
U0001	Sydney CBD	CARROTS PREPACKED 1KG	Feb-16	7000	4000	1.75
U0001	Sydney CBD	CARROTS PREPACKED 1KG	Mar-16	4100	2000	2.05
U0001	Sydney CBD	CARROTS PREPACKED 1KG	Q1-16	16100	8500	1.89
U0002…						

Background

- Transactions data used to compile ~ 25% of CPI

- Stock keeping unit (SKU) defines a product

- Current method directly replaces field collected prices with unit values derived from transactions data within elementary aggregates (Jevons formula)

- Quality benefits: average unit value, increased respondent coverage, informed sampling choices

- Cost benefits: less labour intensive

All groups

Groups

Sub-groups

Expenditure classes

Elementary aggregates
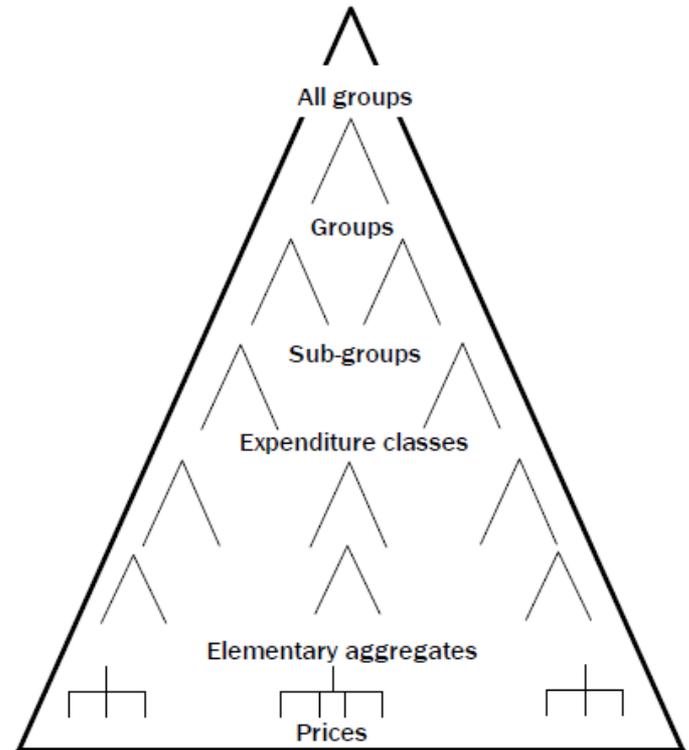
Prices

Background

- While the current method is a significant improvement for the CPI, further enhancements are possible. These enhancements include:

    - Using census of products
    - Weighting prices at the product level
    - Automated processes (less resources)

- What other aggregation options has the ABS considered with transactions data?
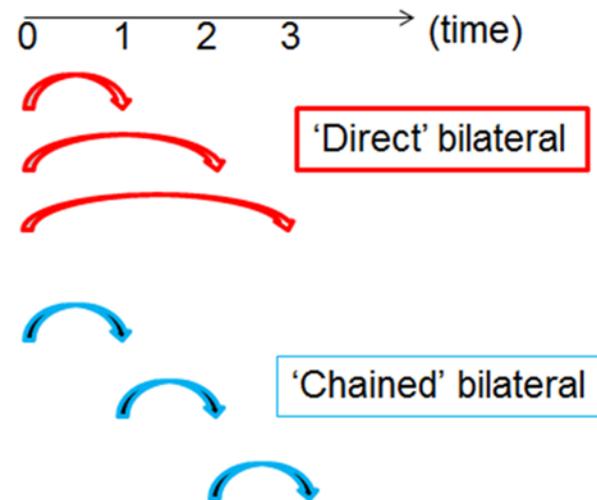
Background

- One option the ABS has considered is a weighted bilateral index formula (e.g. Törnqvist, Fisher)

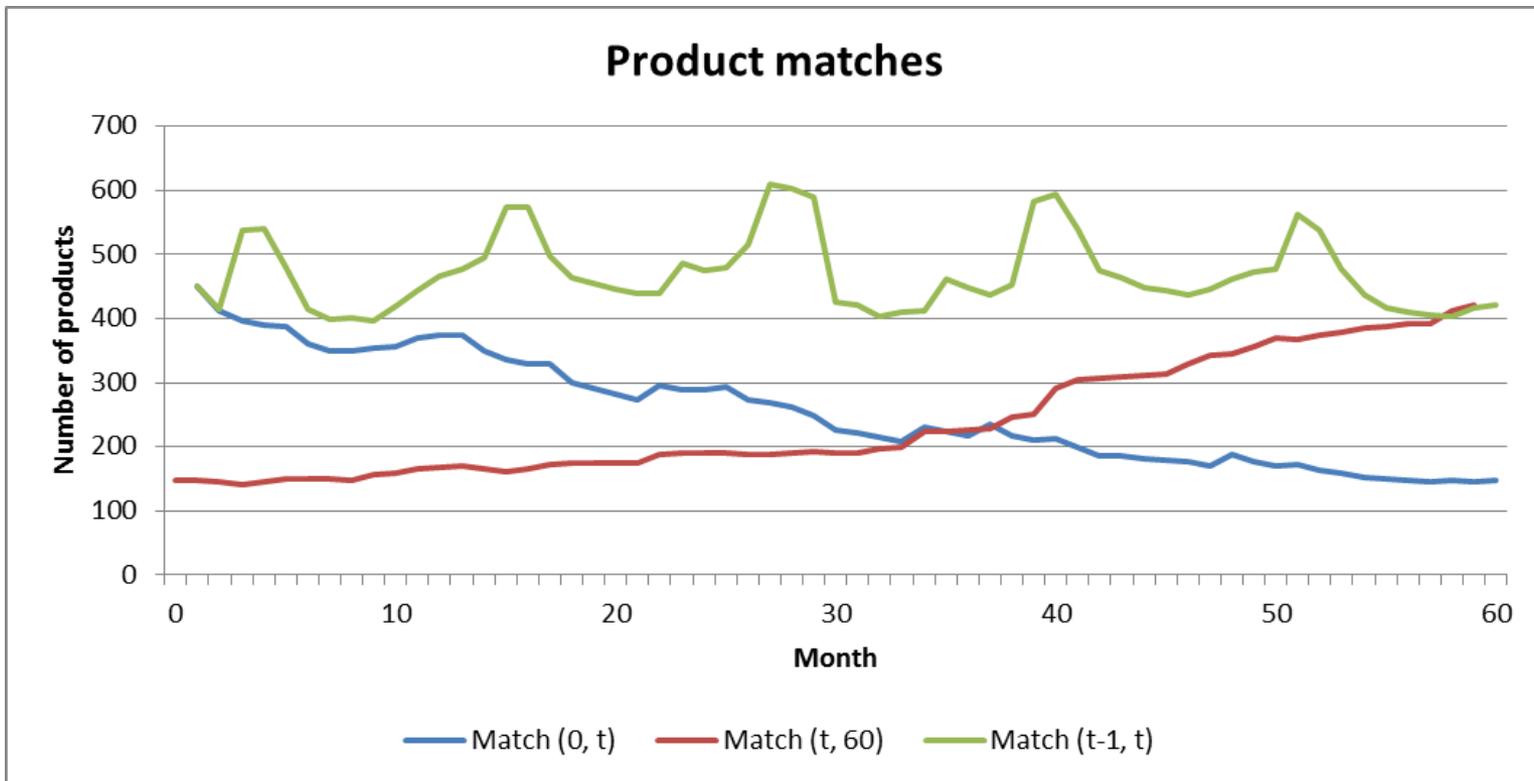- $P\,{}^{t}_{T} = \Pi_i \left(\dfrac{p_i^t}{p_i^0} \right)^{\frac{s_i^t + s_i^0}{2}}$

- $P\,{}^{t}_{F} = \left[P\,{}^{t}_{L} \times P\,{}^{t}_{P} \right]^{\frac{1}{2}}$

- Could use 'direct' or 'chained' weighted bilateral indexes

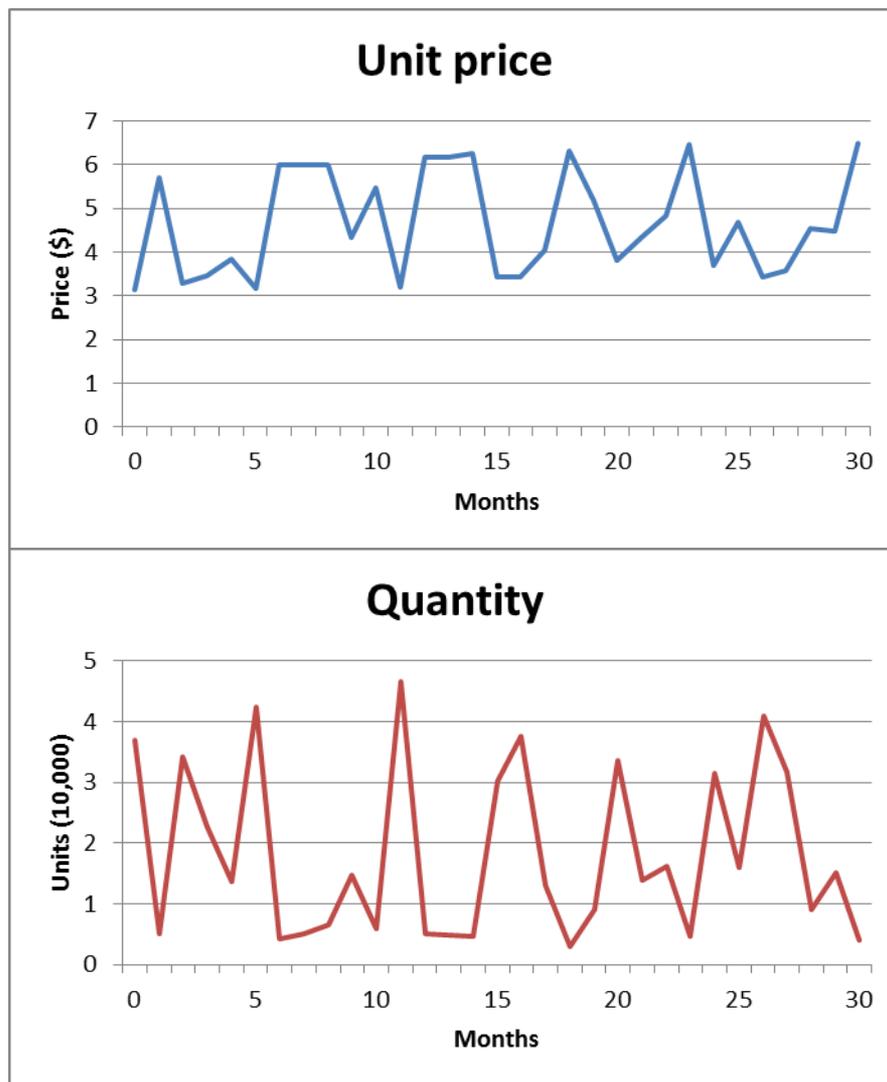- However, dynamic nature of transactions data can make these methods perform poorly

Background

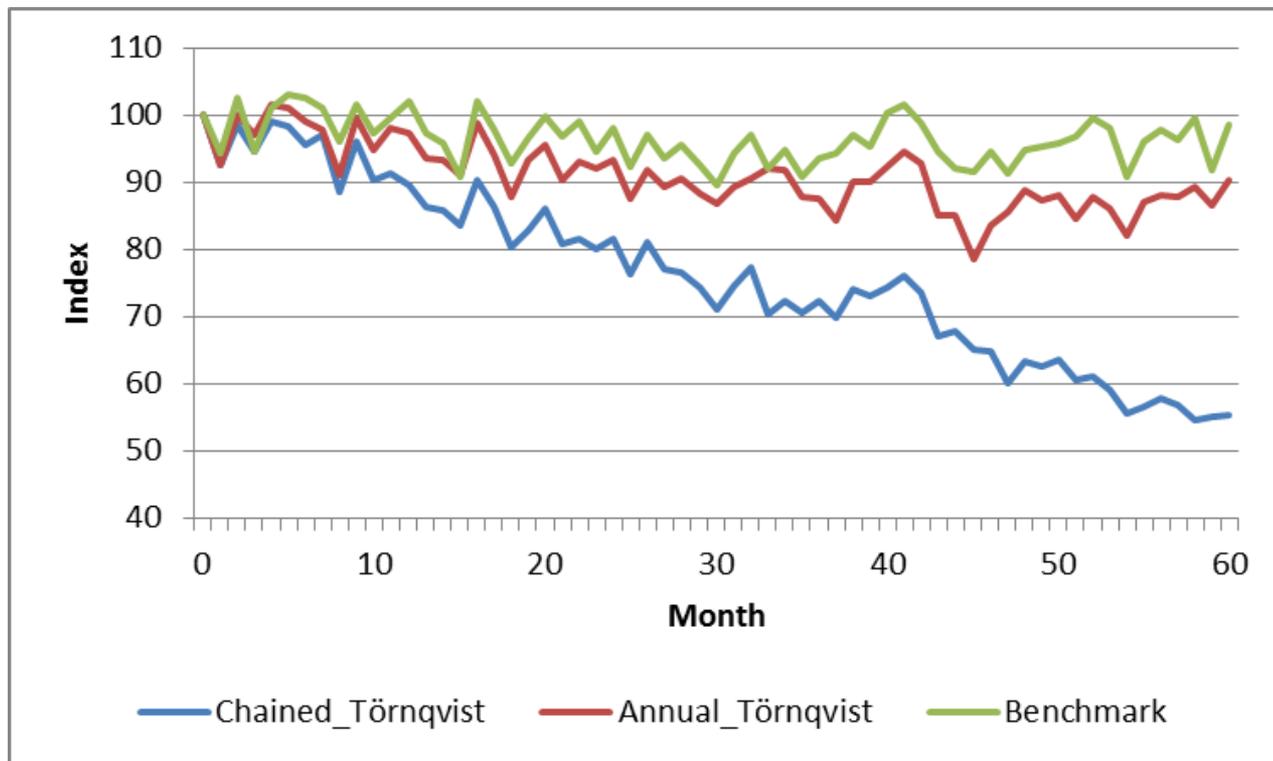- 'Direct' bilateral indexes suffer from a 'matching' problem

Background

- Consumers responsive to sales – price and quantity bouncing can cause problems for chained indexes

Background

- 'Chained' bilateral indexes suffer from a 'chain drift' problem

Background

- Bilateral index methods compare prices across two time periods

- Multilateral index methods make price comparisons across multiple (three or more) time periods

- Historically used in constructing spatial price indexes

- Multilateral methods use all matched products, weight products by economic importance and are free of 'chain drift'
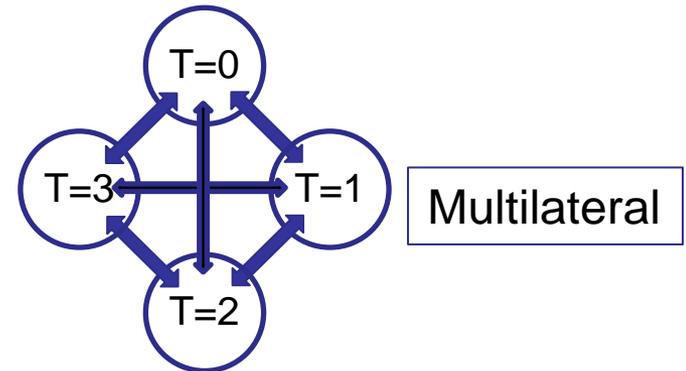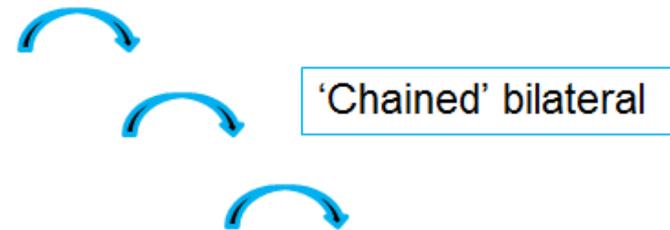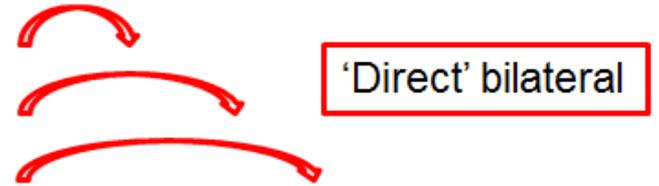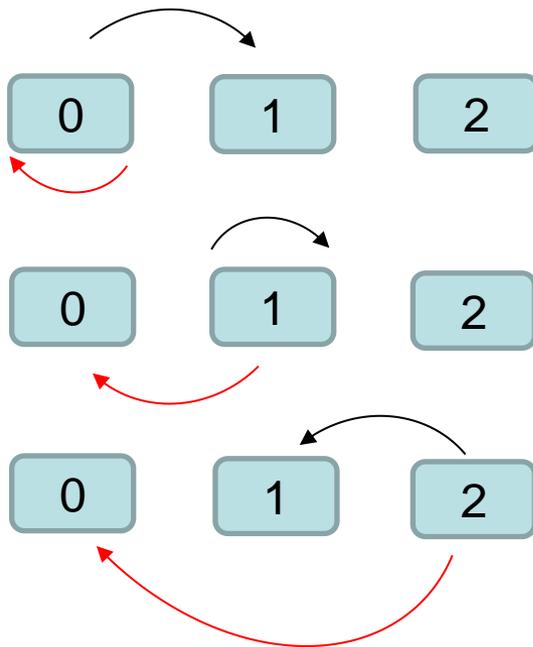
Multilateral method

- ABS has investigated a selection of well-known multilateral methods – this presentation covers two of the most promising methods

1) Gini, Eltetö and Köves, and Szulc (GEKS-Törnqvist): geometric mean of all ratios of bilateral (Törnqvist) indexes where each entity is taken in turn as the base

$$P_{GEKS}^{0,1} = \left(\frac{P^{0,1}}{P^{0,0}} \times \frac{P^{1,1}}{P^{1,0}} \times \frac{P^{2,1}}{P^{2,0}} \right)^{1/3}$$

Multilateral method

2) Time Product Dummy (TPD)

- Similar to hedonic regression models that regress price against a selection of price determining characteristics

- TPD regresses the (log) of price against two dummy variables: product and time

- Weighted least square (WLS) regression using product expenditure share

- Use time dummy parameter as estimate for price change

$$\ln p_i^t = \alpha^0 + \sum_{i=1}^{N-1} \gamma_i D_i + \sum_{t=1}^{T} \delta^t D_i^t + \varepsilon_i^t$$

Extension methods

- When a multilateral method is extended an additional period, previous price movements are revised

- To deal with this revisions problem, the ABS has investigated a selection of extension methods

- These extension methods can be characterised as:

  1. Rolling window approaches (Ivancic, Diewert and Fox 2011, de Haan 2015 , Krsinich 2016, Fox and Diewert 2017)

  2. Direct annual extension (Chessa 2016)

Extension methods

- Rolling window approaches can be joined (spliced) together in different ways (e.g. movement splice, window splice, half splice)

- Rolling window approaches can be various sizes (e.g. 5 quarters, 9 quarters)
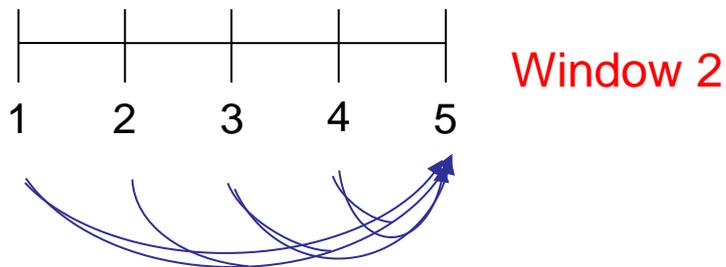
Framework for assessing methods

- The ABS Data Quality Framework used to guide choice of multilateral method

Framework for assessing methods

Criterion	Considerations	Quality dimensions
Resources	Facilitates automation and makes good use of information	*Institutional Environment, Timeliness*
Theoretical properties	Axiomatic and economic approaches to index numbers	*Accuracy*
Transitivity	Risk of drift over time	*Accuracy, Coherence*
Characteristicity	Relevance of bilateral price comparisons to periods at hand	*Accuracy, Relevance*
Flexibility	Scope for adaptation for new products or data sources	*Coherence, Institutional Environment*
Interpretability	Ease of understanding method in general and price movements it calculates	*Interpretability*

Framework for assessing methods

- Using this framework, ABS (2017) recommended the following methods for implementation in the Australian CPI

    - <u>GEKS-Törnqvist</u> as preferred multilateral method

    - Weighted TPD run concurrently (for comparison purposes)

    - Aggregate below the EC level using <u>respondent classes</u>

    - Aggregate respondent classes together using <u>Törnqvist</u> index formula

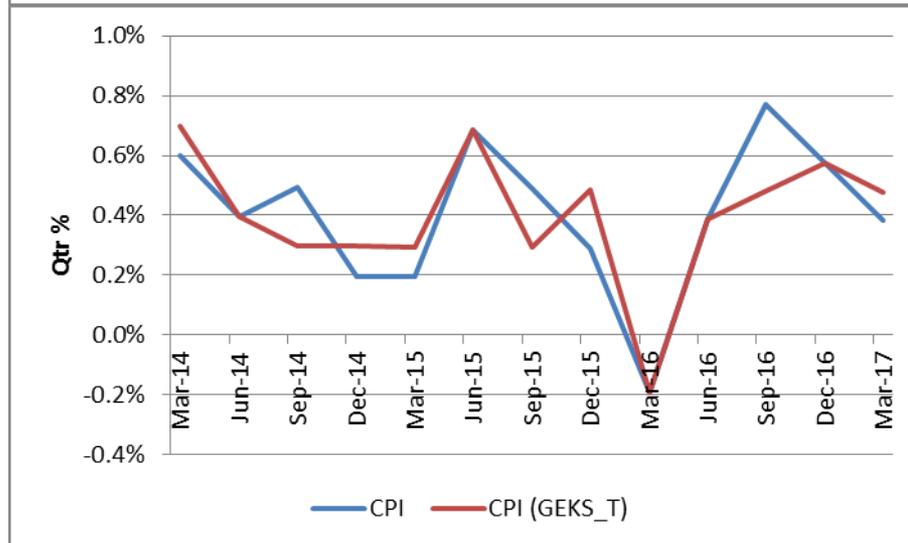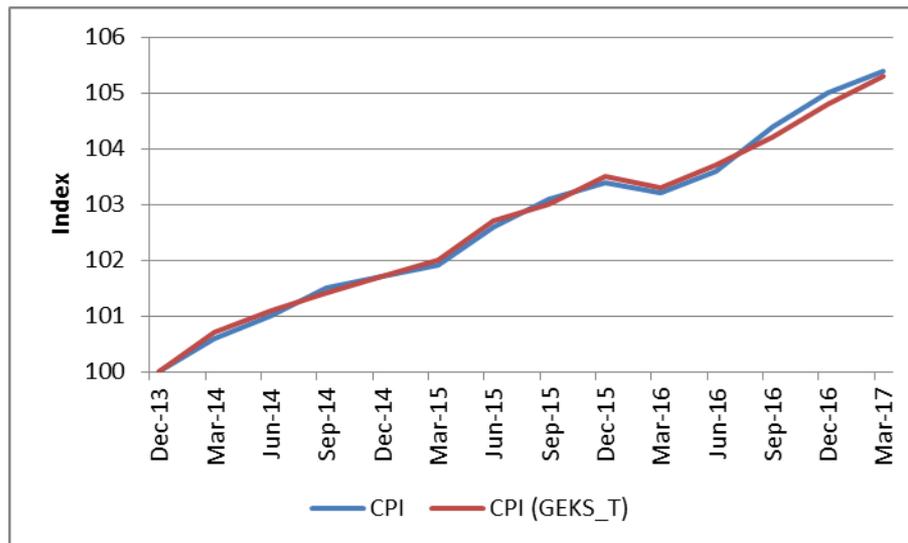   - <u>Mean splice</u> with a rolling window of 9 quarters

Empirical results

- Modified aggregation structure compared for 28 ECs in the CPI

- Respondents weighted by market/expenditure share to produce published level indexes
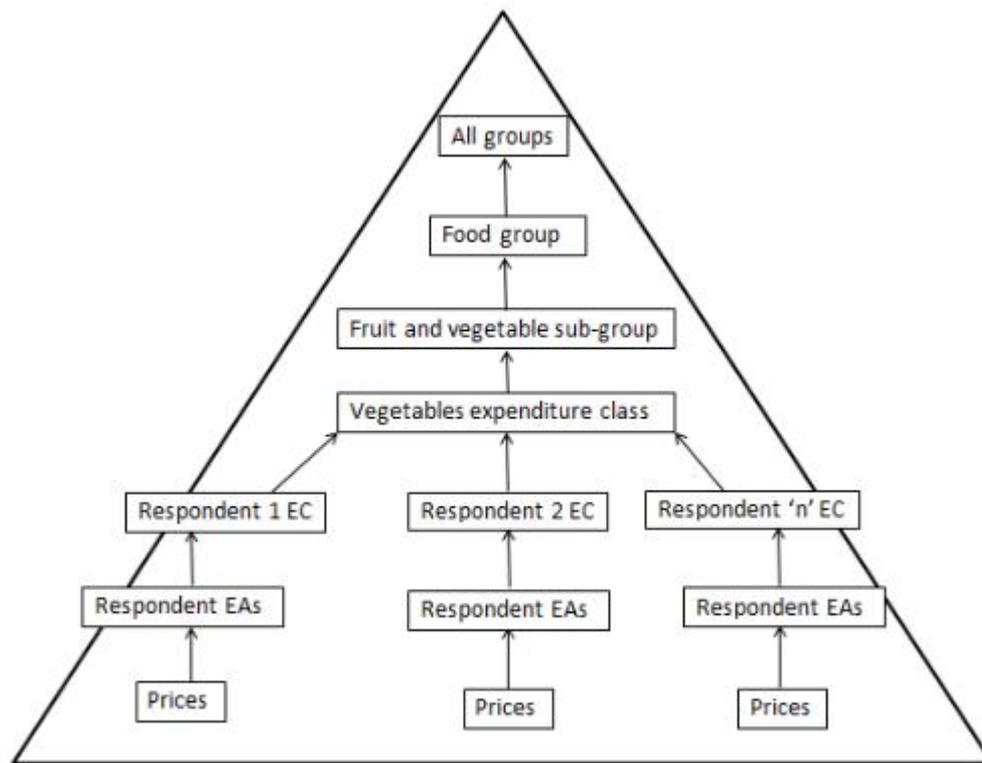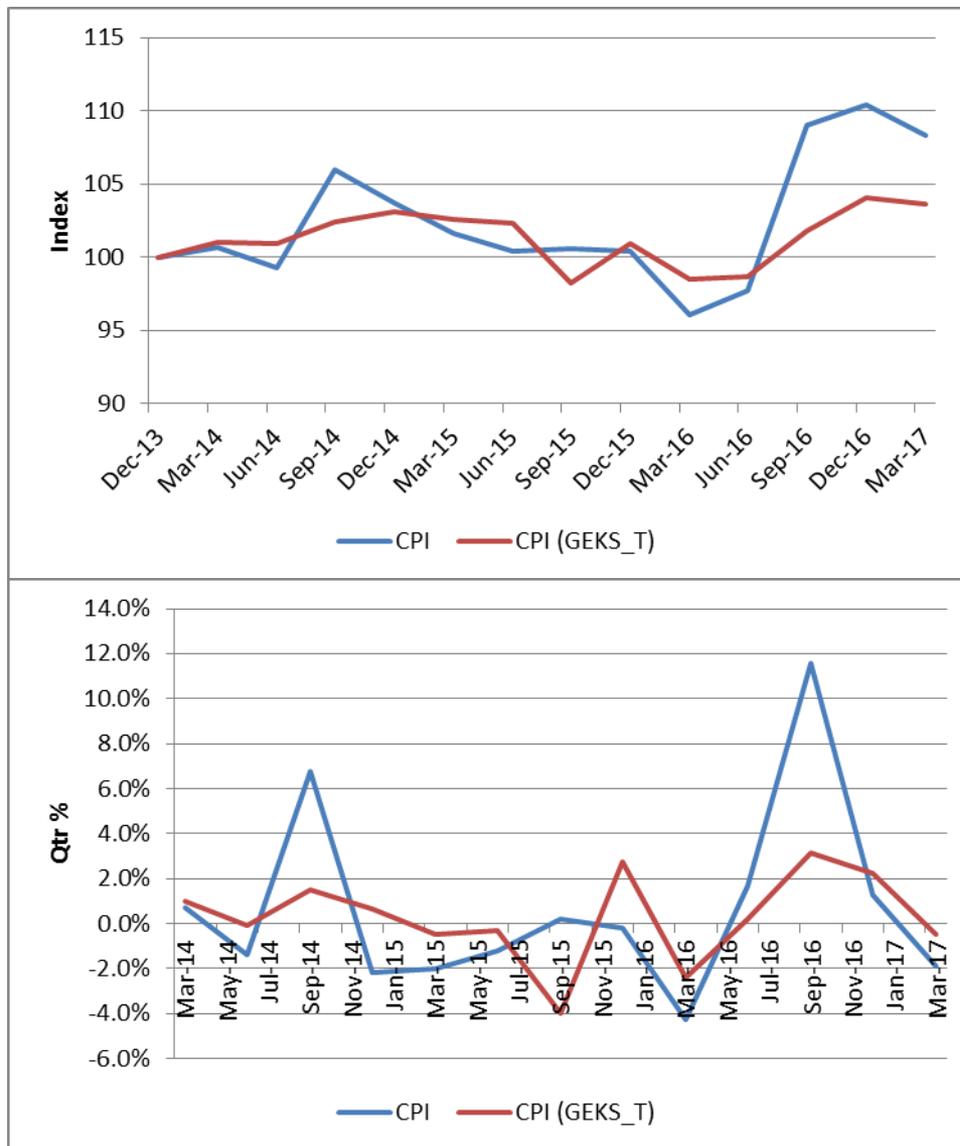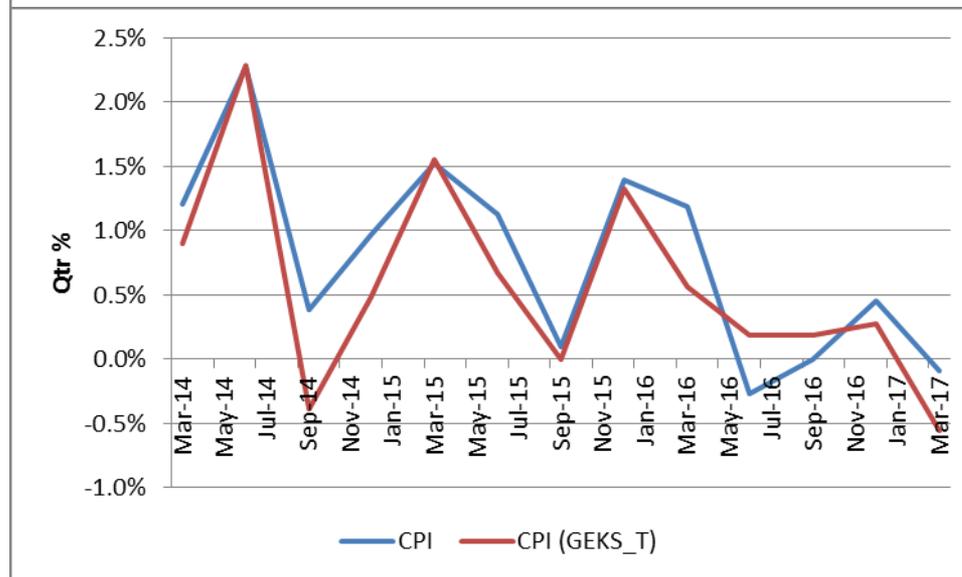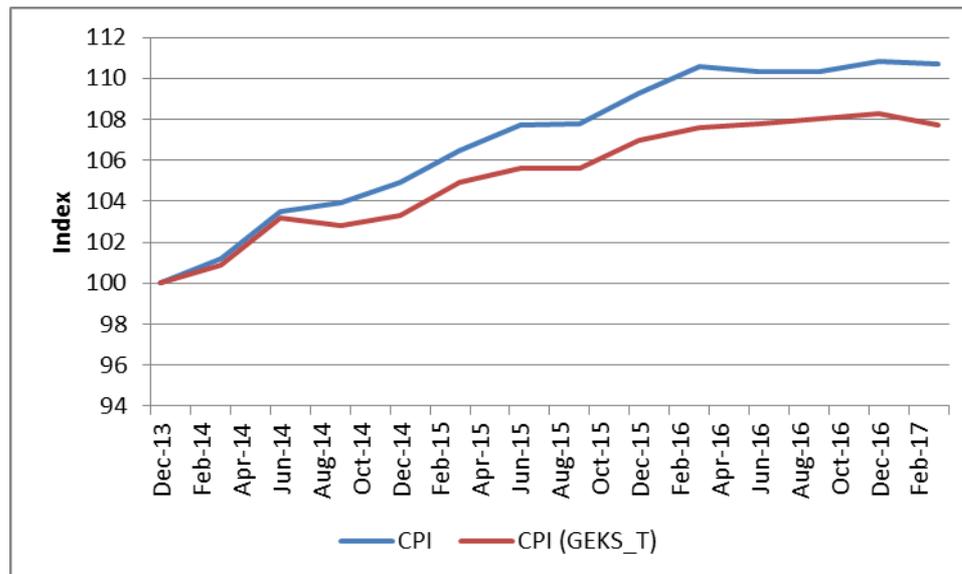
Empirical results

- All Groups CPI

Empirical results
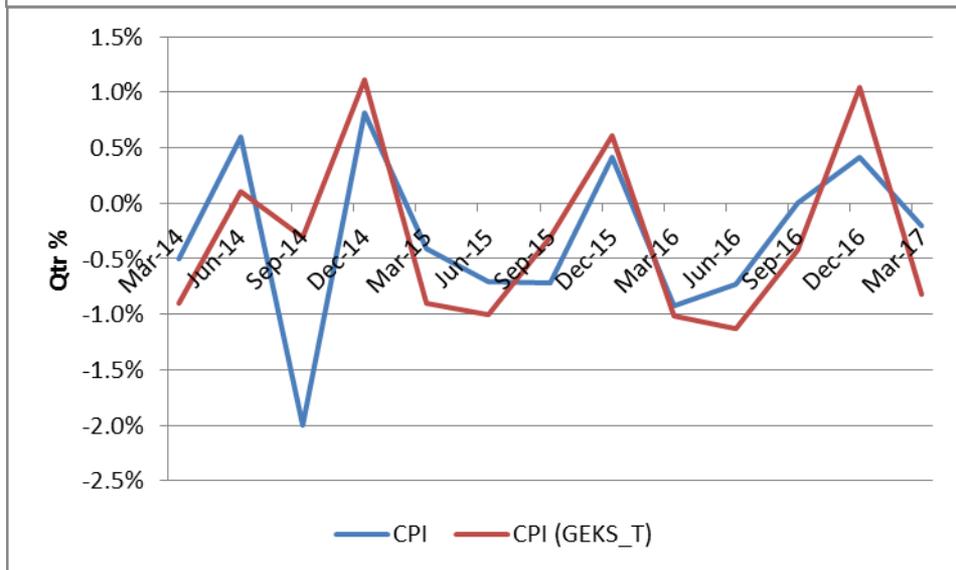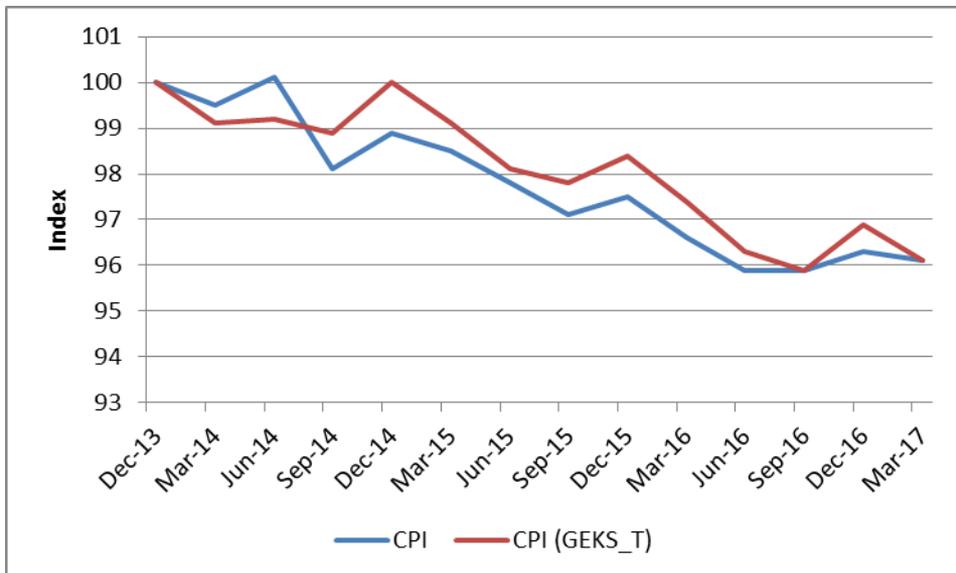
- Fruit and vegetable sub-group

Empirical results

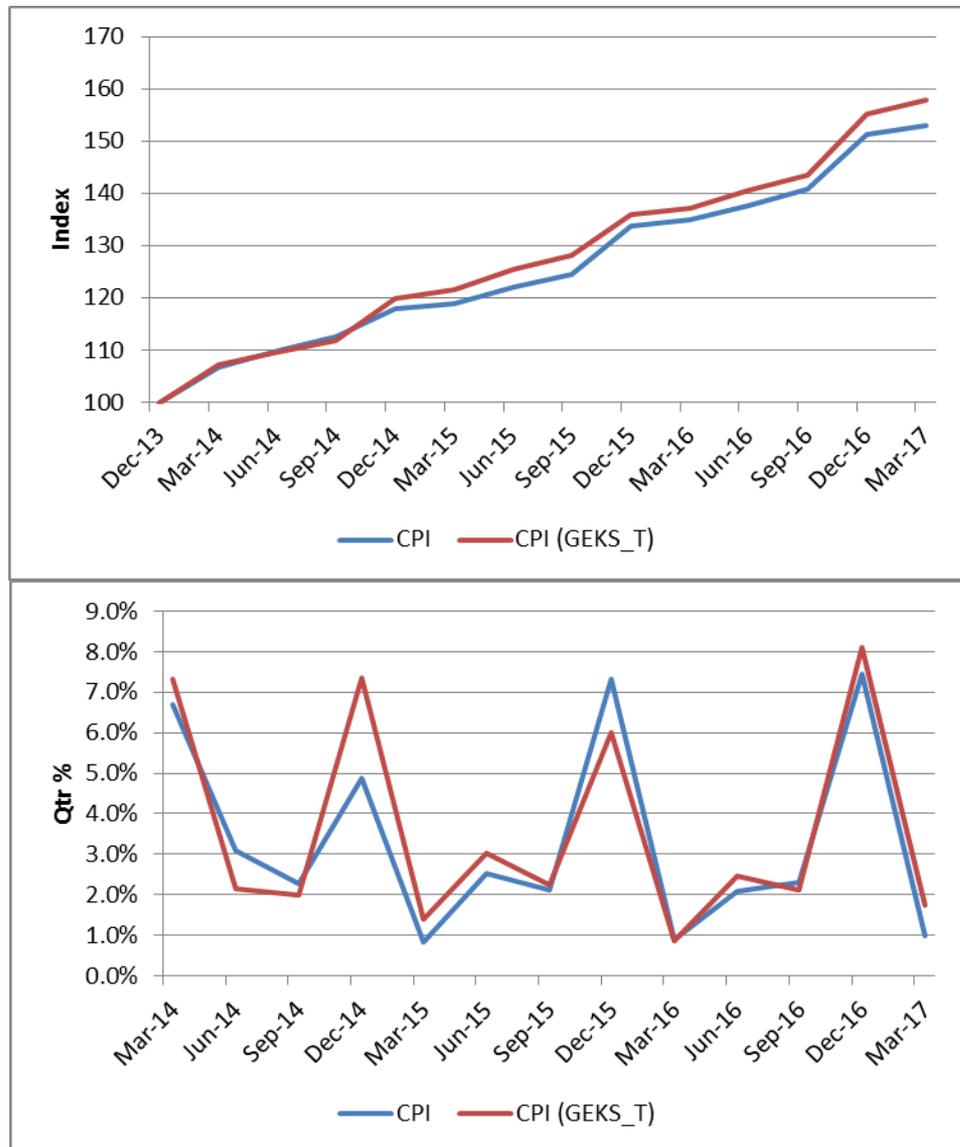- Meat and seafood sub-group

Empirical results

- Bread and cereal products sub-group

Empirical results

- Tobacco sub-group

Consultation

- The ABS has consulted widely with the statistical and user community. Some examples include:

- Release of two information papers (November 2016, June 2017) seeking public submissions

- Peer review of index methods by two international price index experts

- Collaboration with other National Statistical Offices (e.g. New Zealand, Netherlands)

- Bilateral/multilateral workshops with government (e.g. RBA, Treasury)

- Presentations at relevant forums (e.g. Economic Measurement Group @ UNSW, International Working Group on Price Indices)

Implementation

- ABS to continue to monitor these methods during June quarter 2017 and September quarter 2017

- ABS to implement new methods in December quarter 2017 - this is to align the implementation of multilateral methods with the introduction of updated household expenditure data for weighting purposes

References

- Australian Bureau of Statistics (ABS) 2009. ABS Data Quality Framework. cat. no. 1520.0. ABS, Canberra.

- ABS, 2015. Enhancing the Australian CPI: A roadmap. cat. no. 6401.0.60.001. ABS, Canberra.

- ABS, 2016. Information Paper: Making Greater Use of Transactions Data to compile the Consumer Price Index. cat. no. 6401.0.60.003. ABS, Canberra.

- ABS, 2017. Information Paper: An Implementation Plan to Maximise the Use of Transactions Data in the CPI. cat. no. 6401.0.60.004. ABS, Canberra.

- Chessa, A.G. 2016, A New Methodology for Processing Scanner Data in the Dutch CPI, *Eurona* 1/2016, 49-69.

References

- Diewert, W.E. and K.J. Fox 2017, Substitution Bias in Multilateral Methods for CPI Construction using Scanner Data, Discussion Paper 17-02, Vancouver School of Economics, University of British Columbia, Canada.

- de Haan, J. 2015, Rolling Year Time Dummy Indexes and the Choice of Splicing Method, 14th meeting of the Ottawa Group, May 22, Tokyo.

- Ivancic, L., Fox, K. J. & Diewert, E. W. 2011. Scanner data, time aggregation and the construction of price indexes. *Journal of Econometrics,* 161**,** 24-35.

- Krsinich, F. 2016, The FEWS Index: Fixed Effects with a Window Splice, *Journal of Official Statistics* 32, 375-404.