Using Administrative Data for Seasonal Adjustment of Survey Time Series in the Presence of a Major Measurement Change

Xichuan (Mark) Zhang and Nick von Sanden

Australian Bureau of Statistics (ABS)

Outline

- Introduction – what is the problem?

- Literature – what has been done before?

- Our indirect approach - using admin data

- An application for ANZSIC 06

  - Australian and New Zealand Standard Industry Classification 2006

  - Update to the old ANZSIC 93 classification

- Concluding remarks

Introduction

- Major methodological changes in surveys, used to produce time series statistics
  - reduces comparability of data over time
- Inconsistent inter-temporal time series are a major problem in economic research
- Fresh time series measurement after a structural change
  - leads to high risk of misleading signals in Major Economic Indicators (MEIs)

Literature – direct approach

- Bureau of Economic Analysis (1993) explains the impacts of methodological changes on time series
- A "standard" approach to preserving continuity
  - parallel run for the old and new survey methodology for an overlapping transition period
  - the full impact of new survey methodology can then be assessed
  - backcast the old series to be consistent with the new series

Literature – direct approach

- Time series model

$$O_i(t,l) = T_i(t,l) \times S_i(t,l) \times I_i(t,l) \quad , \quad \sum_{l=1}^{P} S_i(t,l) = P$$

where i =1,2 are the old and new survey methodology

$$f:$$
$$\begin{cases} T_2(t,l) = \alpha \times T_1(t,l) \\ S_2(t,l) = \beta(l) \times S_1(t,l), \quad \sum_{l=1}^{P} \beta(l) = P \end{cases}$$

- Time series intervention analysis. eg. Kaiser, R. and Maravall, A. (2001) for both level shift α and seasonal break $\beta(l)$

Literature – direct approach

- Time series intervention analysis requires a parallel run of at least 3 years
- For short time series span
  - McLaren and Zhang (2003) developed a method using a mixed model framework to estimate seasonal structural changes
  - Gatto (2006) developed a "heuristic" method using a regression framework

Our indirect approach

- Indirect approach via administrative data source

old   new

survey   (1,1) ← $f(1,2:1,1)$ (1,2)

$f(1,1:2,1)$ $f(1,2:2,2)$

admin   (2,1) ← $f(2,1:2,2)$ (2,2)
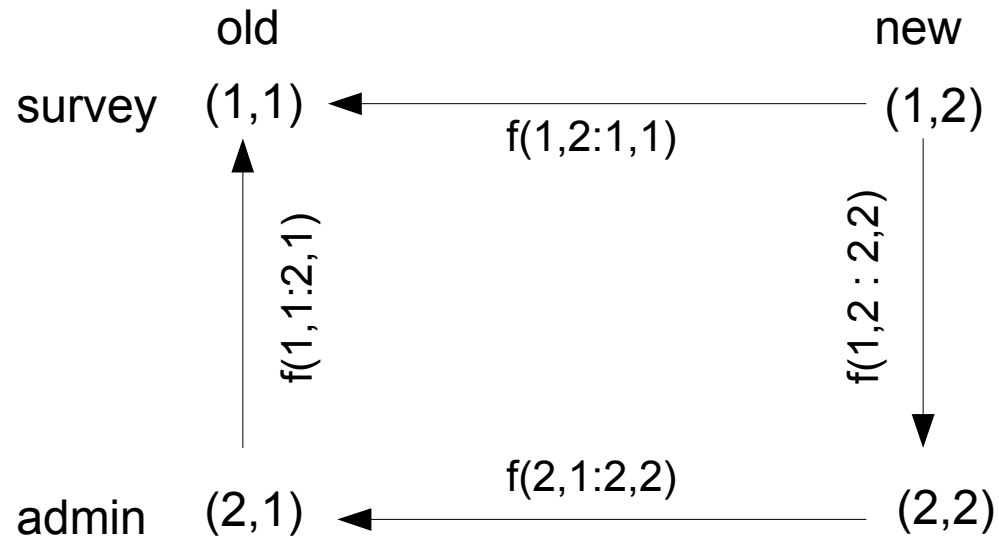
Proposed approach

- **STEP 1**: conceptually similar time series data is sourced from an administrative data set and the seasonal pattern change from the old to the new methodology is estimated based on this administrative time series

- **STEP 2**: a short parallel run is conducted (e.g. 2~3 periods) on the survey time series and the seasonal pattern change estimated from STEP 1 is applied. This allows us to produce a more appropriate estimate of the level shift from the old to the new method on the survey data

An application

- Background:
  - The ABS is preparing to implement ANZSIC 06 for all ABS surveys and the National Accounts
    - by September quarter 2009
    - replaces existing ANZSIC 93
  - Earlier strategy: 1 year plus 1 period parallel run
 - New challenge: parallel run will be conducted for only 2 periods

An application - QBIS

- ABS has run Quarterly Business Indicators Survey (QBIS) since 2001
- Provides 22 ANZSIC 93 divisions and subdivisions
  - sales of goods and services
  - **wages and salaries**
  - profits, and
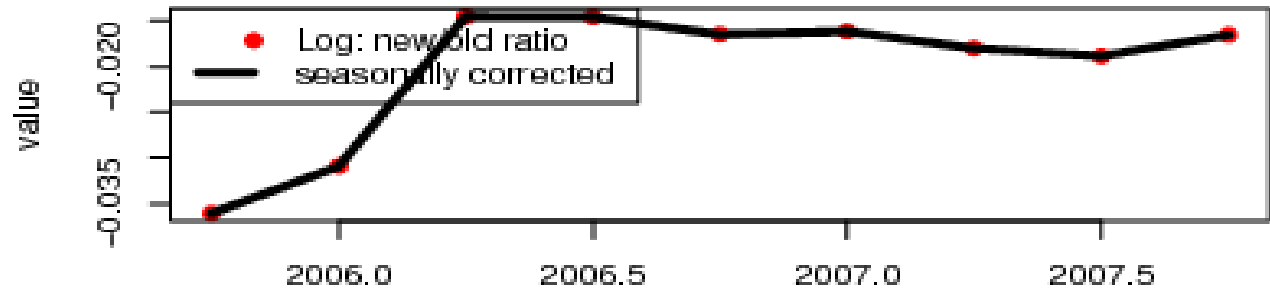 - inventories

An application - QBIS

- Business activity statement Unit Report Estimates (BURE)
  - administrative data from Australian Taxation Office (ATO)
- Contains information on
  - business turnover
  - **wages**
  - capital expenditure
  - non-capital expenditure
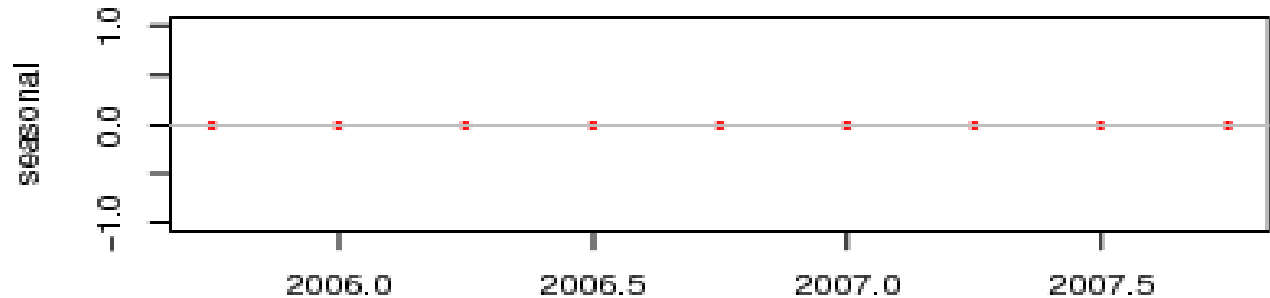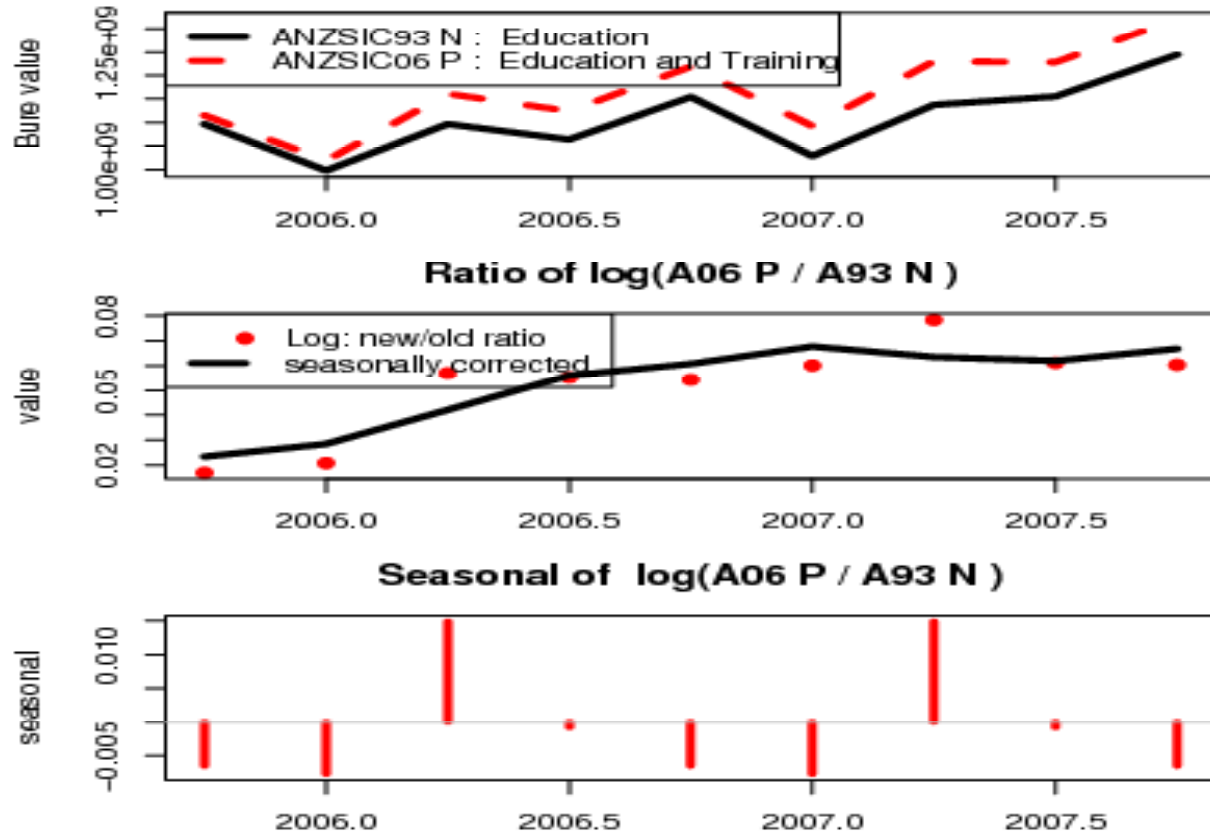- coded in both ANZSIC 93 (since 2002), and ANZSIC 06 (since 2005)

QBIS Wages
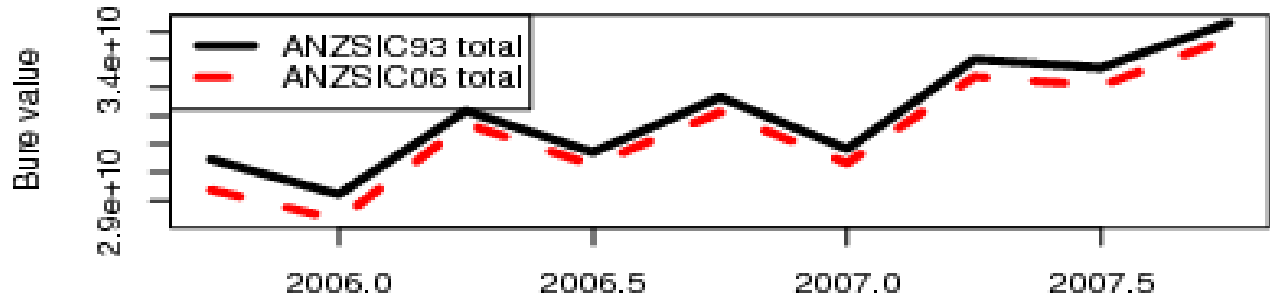
QBIS Wages – Education

Reason: ANZSIC 06 - Educational Support Services

Concluding remarks

- the proposed method is a feasible solution under the scenario that long periods of overlapping parallel runs are not affordable

  – by using administrative data sources

- the proposed method works under a very plausible assumption

  – seasonal impacts are the same on both surveys and administrative data

- the proposed method can be used on certain higher classification levels

- a survey item may not have an equivalent in the administrative data source

Thanks
&
Questions