Implementation of consumer electronics scanner data in the New Zealand CPI

Frances Krsinich, Statistics New Zealand

Japan, May 2015

Ottawa Group 2015

Introduction

- Consumer electronics scanner data from GfK
- 12 product categories
- ITRYGEKS method
- Implemented in NZ CPI September 2014 quarter

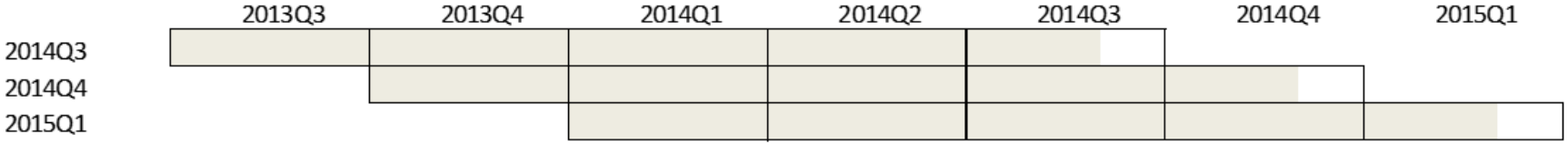
OBS #	PERIOD	QUANTITY SOLD	TOTAL SALES	MODEL	BRAND	CF CARD	CHIPTYPE	DIGITAL INDEX
1	Mar-11	7	2277.59	*	*	NO CF CARD	CCD	DIGITAL INDEX
2	Mar-11	388	146213.38	*	*	N.A.	CCD	DIGITAL INDEX
3	Mar-11	152	67078.15	*	*	N.A.	CCD	N.A.
4	Mar-11	143	88634.93	*	*	NO CF CARD	CCD	DIGITAL INDEX
5	Mar-11	132	80531.68	*	*	NO CF CARD	CCD	DIGITAL INDEX
6	Mar-11	103	109032.06	*	*	NO CF CARD	CMOS	DIGITAL INDEX
7	Mar-11	83	27314.77	*	*	NO CF CARD	CCD	DIGITAL INDEX
8	Mar-11	58	53096.64	*	*	NO CF CARD	CMOS	DIGITAL INDEX
9	Mar-11	43	40971.79	*	*	N.A.	CMOS	N.A.
10	Apr-11	43	29163.53	*	*	NO CF CARD	CCD	N.A.
11	Apr-11	37	43785.84	*	*	NO CF CARD	CMOS	DIGITAL INDEX
12	Apr-11	29	24746.86	*	*	NO CF CARD	CMOS	N.A.
13	Apr-11	27	11563.74	*	*	NO CF CARD	CCD	DIGITAL INDEX
14	Apr-11	25	10852.06	*	*	NO CF CARD	CCD	DIGITAL INDEX
15	Apr-11	23	17257.10	*	*	NO CF CARD	CMOS	DIGITAL INDEX
16	Apr-11	17	25483.29	*	*	NO CF CARD	CMOS	DIGITAL INDEX
17	Apr-11	14	13488.18	*	*	NO CF CARD	CMOS	N.A.
18	Apr-11	13	25640.43	*	*	NO CF CARD	CMOS	N.A.
19	Apr-11	17	18819.06	*	*	NO CF CARD	CMOS	DIGITAL INDEX
20	Apr-11	16	6821.91	*	*	NO CF CARD	CMOS	N.A.

OBS #	DIGITAL INPUT	HD Formats	IMAGE STABIL.	LCD SCREEN SIZE	MEMORY CAPACITY	OPTICAL ZOOM	OUTDOOR FUNCTIO	PIXEL TOTAL
1	NO DIG. INPUT	SD	ELEC.STAB	2.7	60	32	N.A.	1.07
2	NO DIG. INPUT	SD	OPTICAL STAB.	2.7	N.A.	70	NO WATER_SHOCK	0.8
3	DIGITAL INPUT	SD	ELEC.STAB	2.7	4	60	NO WATER_SHOCK	1.07
4	NO DIG. INPUT	SD	OPTICAL STAB.	2.7	80	78	N.A.	0.8
5	DIGITAL INPUT	SD	ELEC.STAB	2.7	80	60	NO WATER_SHOCK	0.68
6	NO DIG. INPUT	HD HDD	OPTICAL STAB.	2.7	120	25	NO WATER_SHOCK	3.32
7	DIGITAL INPUT	SD	ELEC.STAB	2.7	N.A.	39	NO WATER_SHOCK	0.8
8	DIGITAL INPUT	HD MEMORY	OPTICAL STAB.	2.7	N.A.	25	N.A.	3.32
9	DIGITAL INPUT	HD MEMORY	ELEC.STAB	2.7	8	25	N.A.	2.36
10	NO DIG. INPUT	SD	ELEC.STAB	2.7	16	37	N.A.	0.8
11	DIGITAL INPUT	HD HDD	ELEC.STAB	2.7	120	25	NO WATER_SHOCK	2.36
12	NO DIG. INPUT	HD MEMORY	ELEC.STAB	3	4	10	N.A.	10
13	NO DIG. INPUT	SD	ELEC.STAB	2.7	N.A.	52	NO WATER_SHOCK	0.8
14	NO DIG. INPUT	SD	ELEC.STAB	2.7	N.A.	37	NO WATER_SHOCK	0.8
15	NO DIG. INPUT	HD MEMORY	OPTICAL STAB.	2.7	N.A.	20	NO WATER_SHOCK	3.32
16	NO DIG. INPUT	HD MEMORY	OPTICAL STAB.	3	N.A.	12	N.A.	9.15
17	NO DIG. INPUT	HD MEMORY	ELEC.STAB	2.7	32	20	N.A.	2.39
18	NO DIG. INPUT	HD MEMORY	OPTICAL STAB.	3.5	64	10	NO WATER_SHOCK	6.63
19	DIGITAL INPUT	HD HDD	OPTICAL STAB.	2.7	160	12	NO WATER_SHOCK	7.1
20	NO DIG. INPUT	HD MEMORY	ELEC.STAB	2.7	8	30	NO WATER_SHOCK	3.32

Product category	Characteristics	Products
Heat pumps	27	72
Desktop computers	56	107
Laptop computers	71	445
Tablet computers	73	148
Multi-function devices	53	102
Cellphone handsets	59	392
Digital cameras	77	228
Digital camera memory cards	10	254
Television sets	62	325
Set-top boxes for television sets	49	24
DVD, Blu-ray players, and player/recorders	50	129
Home theatre and stereo systems	62	224

Practical issues

- Incomplete data in time for production



- Monthly data – quarterly indexes

Methodology

- Issues with price measurement from scanner data:
  - Product turnover (particularly high for consumer electronics)
  - Price/quantity volatility – chain drift

- Full set of characteristics in GfK data so we can use the ITRYGEKS index

ITRYGEKS

- Imputation Törnqvist rolling year GEKS (de Haan and Krsinich, 2014)
- Extension of RYGEKS (Ivancic, Diewert & Fox, 2011) based on bilateral time-dummy hedonic indexes (rather than superlative indexes eg Fisher, Törnqvist)
- Rolling window, latest movement spliced on, so index is unrevisable

ITRYGEKS

- RYGEKS:

$$P_{RYGEKS}^{0,13} = P_{GEKS}^{0,12} \prod_{t=1}^{13} \left(P^{12,t} \times P^{t,13} \right)^{1/13} = \prod_{t=0}^{12} \left(P^{ot} \times P^{t,12} \right)^{1/13} \prod_{t=1}^{13} \left(P^{12,t} \times P^{t,13} \right)^{1/13}$$

- ITRYGEKS: replace the superlative indexes *P* with bilateral time-dummy hedonic indexes

$$P_{ITRYGEKS}^{0t} = \prod_{i \in U^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{\frac{s_i^0 + s_i^t}{2}} \prod_{i \in U_D^{0t}} \left(\frac{\hat{p}_i^t}{p_i^0} \right)^{\frac{s_i^0}{2}} \prod_{i \in U_N^{0t}} \left(\frac{p_i^t}{\hat{p}_i^0} \right)^{\frac{s_i^t}{2}}$$

The processing system

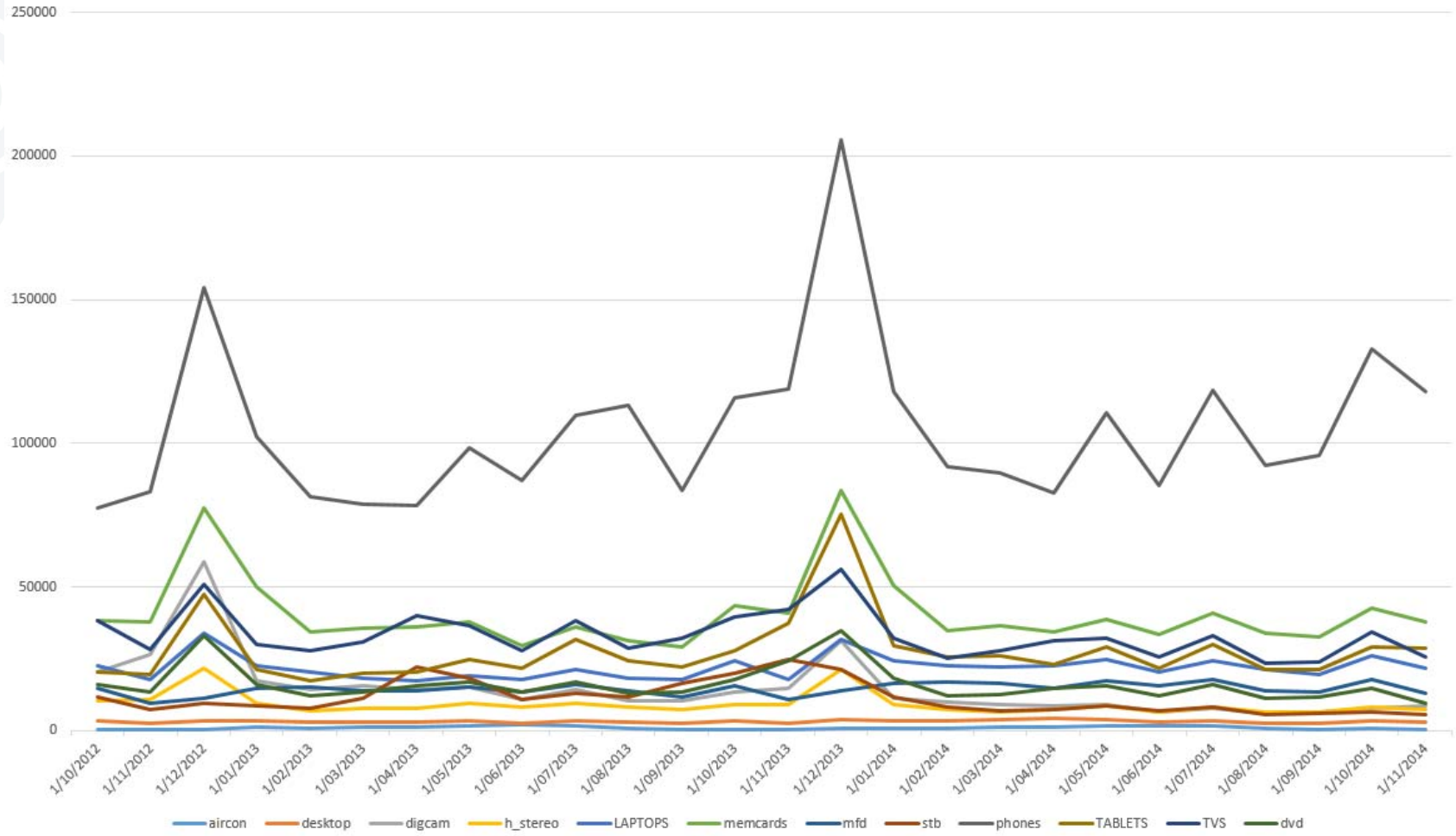
- User-written prototype in SAS
- Iterative (new issues emerge)
- Will inform and become part of a more general Prices big data system:
  - supermarket scanner data
  - online data
  - customs data
  - telecommunications bills data
  - survey data (eg rents, used cars) with hedonic index estimation
  - other as-yet unspecified ‘big data’

Monitoring and analysis

- Processes include:
  - Checking characteristics haven't changed
  - Time series of quantities sold
  - Time series of (unadjusted) average prices
  - Outlier identification (movements, levels, longitudinal record)
  - Distribution across key characteristics (quantities and expenditure shares)
  - Monthly ITRYGEKS indexes
  - Statistics on product turnover

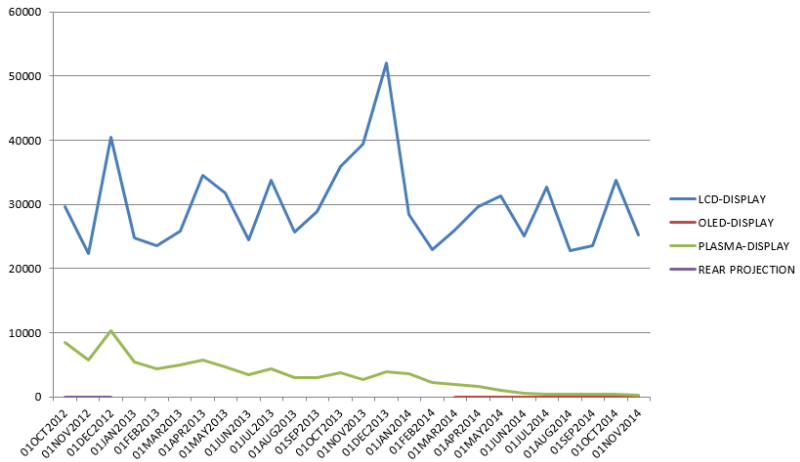
- If necessary we can:
  - remove outliers and rerun (so far, this makes very little difference)
  - run indexes on subsets of data (eg brands)

Quantities

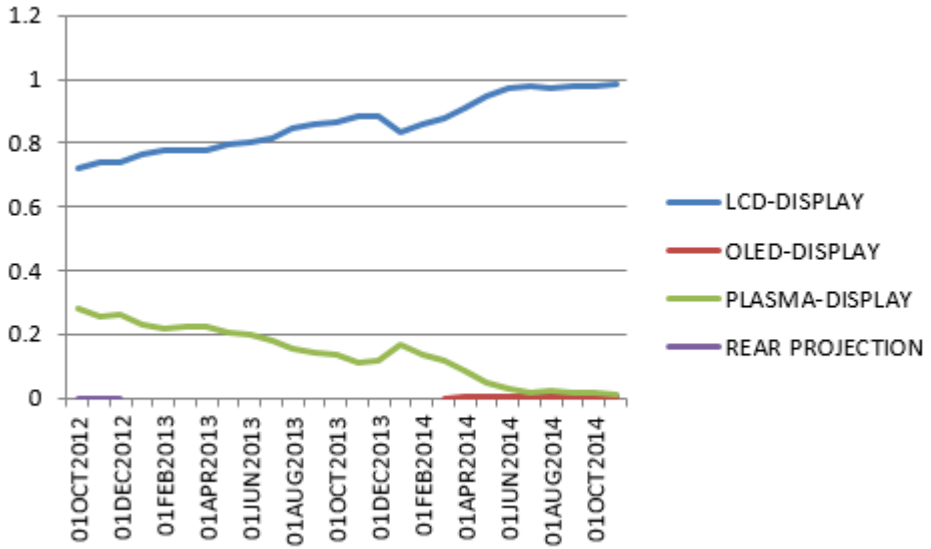

250000
200000
150000
100000
50000
0
1/10/2012
1/11/2012
1/12/2012
1/01/2013
1/02/2013
1/03/2013
1/04/2013
1/05/2013
1/06/2013
1/07/2013
1/08/2013
1/09/2013
1/10/2013
1/11/2013
1/12/2013
1/01/2014
1/02/2014
1/03/2014
1/04/2014
1/05/2014
1/06/2014
1/07/2014
1/08/2014
1/09/2014
1/10/2014
1/11/2014
aircon
desktop
digcam
h_stereo
LAPTOPS
memcards
mfd
stb
phones
TABLETS
TVS
dvd

Distributions of characteristics

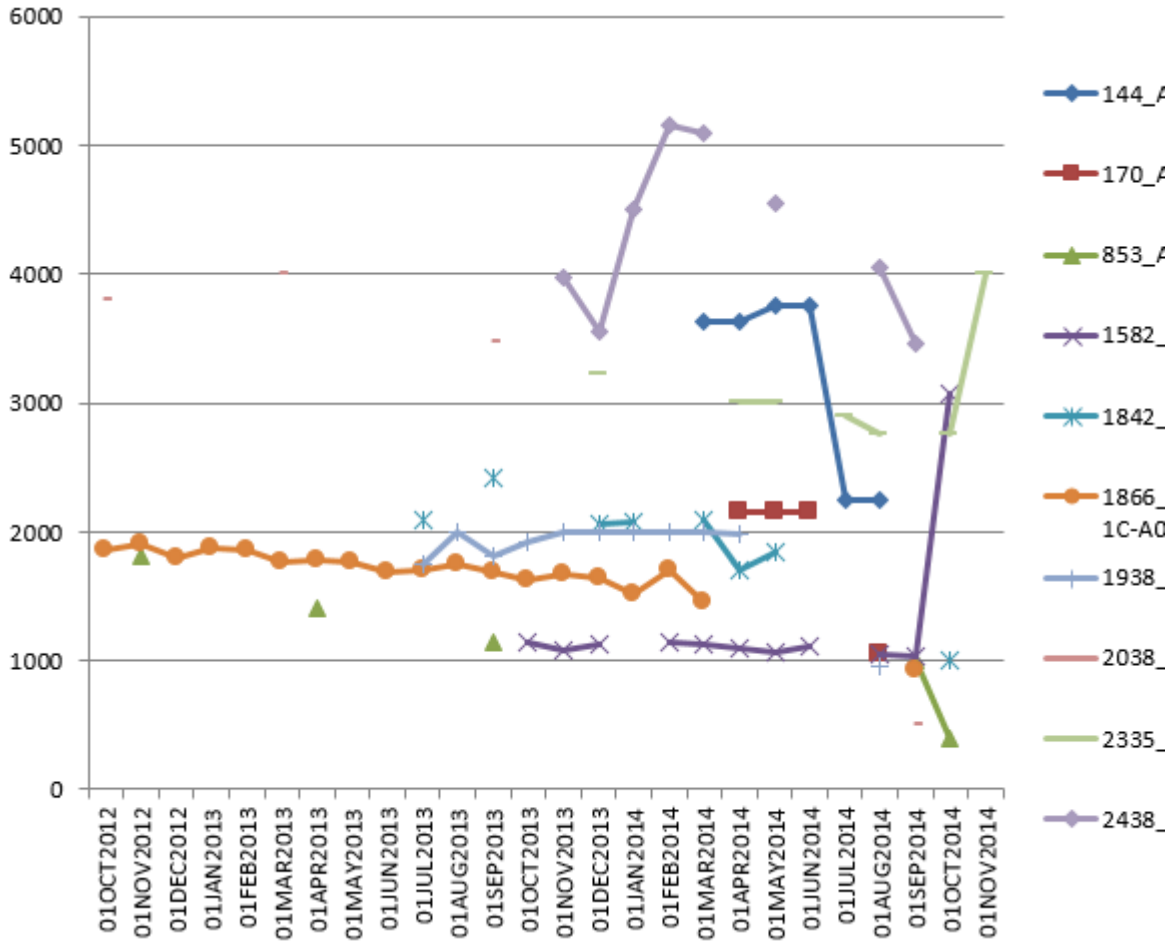
Quantities



Expenditure shares



Outliers


6000
5000
4000
3000
2000
1000
0
01OCT2012
01NOV2012
01DEC2012
01JAN2013
01FEB2013
01MAR2013
01APR2013
01MAY2013
01JUN2013
01JUL2013
01AUG2013
01SEP2013
01OCT2013
01NOV2013
01DEC2013
01JAN2014
01FEB2014
01MAR2014
01APR2014
01MAY2014
01JUN2014
01JUL2014
01AUG2014
01SEP2014
01OCT2014
01NOV2014
144_A
170_A
853_A
1582_
1842_
1866_
1C-A0
1938_
2038_
2335_
2438_

Communicating the new development

- Variety of different forums
- Drill-down from ‘plain English’ to technical explanations in ‘sources and methods’ doc on website released with first quarter
- Price Index News (quarterly) – signaled upcoming change and linked to info afterwards
- Newsletters:
  - NZ Association of Economist newsletter ‘*Asymmetric Information*’
  - IASS newsletter ‘*Survey Statistician*’
  - SNZ ‘Expert Data Users’ newsletter

Conclusion

- Consumer electronics scanner data introduced in NZ CPI September 2014
- ITRYGEKS index
- Iterative approach to developing processes for monitoring, analysis & estimation
- First step of a more general Prices big-data system in Statistics New Zealand
- Variety of communications to users