Consumer Price Index Biases

-Elementary Index Biases vs. Sampling Biases-

Satoshi Imai
(Statistics Bureau of Japan)

Erwin Diewert
(University of British Columbia and New South Wales University)

Chihiro Shimizu
(National University of Singapore and University of British Columbia)

Outline and Motivation

- **Consumer price indices (CPIs) in particular are one of the indices in economic statistics that is most closely linked with policy.**
- Many policy debates have arisen surrounding price index biases.

- **The Biases caused by research design, technical limitations and budget constraints**
- It is extremely important to provide index users with some indication about the magnitude of possible biases.
- Clarifying the factors causing such biases, it will be easier to make future improvements.

Satoshi Imai s2.imai@soumu.go.jp

Outline and Motivation

- **The Biases**
- This paper divided possible CPI biases into
  biases based on differences in the calculation method
  (elementary index biases) and
  biases based on sampling (sampling method biases),
  with the aim of clarifying the magnitude of each type of bias.

- **The purpose of this paper is to observe how changing the sampling method and the choice of elementary index formula changes the resulting price index.**
- This paper do not provide estimates of "biases" in actual Consumer Price Indexes.
- But our results may be helpful in improving actual price indexes in the future.

Satoshi Imai s2.imai@soumu.go.jp

Outline and Motivation

- **The Elementary Index Biases**
- Many statistical agencies measure price indices based on the Laspeyres formula (in practice, the Lowe formula).
- However, besides this method, many other formulae have been proposed depending on the index's purpose, such as Carli, Dutot, Jevons, weighted Jevons, Törnqvist, Lowe, Paasche, and Fisher indices.
- In theory, Laspeyres-based indices have an upward bias and Paasche-based indices have a downward bias.
- we use to calculate the item level indices are Carli, Dutot, Jevons, weighted Jevons, Törnqvist, Lowe (annual weight), Laspeyres, Paasche, and Fisher.
- We use to aggregate item level indices into national indices are Lowe (annual weight), Laspeyres.
- In order to conduct comparison at the aggregation stage, we calculated a two-stage Fisher index, too.

Satoshi Imai s2.imai@soumu.go.jp

Outline and Motivation

- **The Sampling Methods Biases**
- With regard to problems based on purposive sampling, significant differences still exist at the item level, but it is said that these may be ignored at upper levels of aggregation.
- These discussions focused mainly on biases using purposive sampling vs probabilistic random sampling.
- But in practice, we have another important choice for sampling, i.e. the sampling size and “representativeness”.

- The sampling size relates to how much outlets are selected for each items.
- The representativeness relates to how much products are selected for given outlets for each items.

Satoshi Imai s2.imai@soumu.go.jp

Outline and Motivation

- **The Sampling Methods Biases (the sampling size)**
- With regard to the selection of outlets, in order to compare outlets with a large share of the price data used at a given month t to those with a small share, we sampled outlets according to the size of their monthly customer base.
- This was done in order to look at the effect of selecting more low-popularity outlets for month t by expanding the number of outlets to be used in creating the index.
- That is, it is possible that low-popularity outlets may adopt different pricing strategies from high-popularity outlets, and we wanted to examine the effect of including such outlets in the survey.

Outline and Motivation

- **The Sampling Methods Biases (the representativeness)**
- With regard to selecting products at the target outlets sampled based on the above procedure, we used two methods:
  1) selecting only the top-ranked product in terms of sales volume
  2) selecting the top five products.

- This was done because we anticipated that the pricing strategy for the most salable product would differ from the pricing strategy for less salable products.

- It may be expected that biases due to the price sampling method will be canceled out when aggregated at the upper level.

Satoshi Imai s2.imai@soumu.go.jp

Sampling Data Number of Outlets, Products and Observations

	No. of outlets	No. of products	No. of observations
2000	100	3,652	4,010,804
2001	100	3,911	4,272,381
2002	100	3,790	4,408,992
2003	100	3,851	3,987,877
2004	100	3,951	4,752,444
2005	100	3,832	4,839,600
2006	100	4,029	4,981,847
2007	100	4,139	5,093,593
2008	100	4,369	5,089,374
2009	100	4,220	5,075,707
2010	100	4,246	5,257,439
2011	100	4,592	5,272,547
2012	100	4,630	5,677,755
2013	100	4,943	5,780,280
2014 *	100	4,523	3,369,542

* Jan. 2014 – Jul. 2014 only

We used Nikkei POS Dataset from January 2000 to July 2014.

Satoshi Imai s2.imai@soumu.go.jp

Categories

Code	Category Name
234001	Non-glutinous rice
137001	Cup instant noodle
031031	Bacon
046001	Fresh milk (1L)
041001	Butter
212001	Canned beer
601021	Shampoo (refill)
610001	Facial tissue (boxed)
612042	Liquid detergent for general clothing (refill)
818001	Dog food (dry type)
[Seasonal one]	
191001	Chocolate bar
107008	Dried Japanese vermicelli
620001	Clothing insect repellent

These 13 product categories account for about 2.1% of the applicable CPI’s weight.

Satoshi Imai s2.imai@soumu.go.jp

Sampling Classes

	No. of outlets (sampling size)	No of items from each outlet (representativeness)	Monthly observation
Sampling A	10	Top 5	50
Sampling B	20	Top 5	100
Sampling C	50	Top 1	50
Sampling D	100	Top 1	100

- Sampling Methods A and B, this was done in order to reflect a broader range of product substitution in the index calculations by using weights corresponding to the sales volume in the available data.

- Sampling Methods C and D, this simulates cases where existing national statistics offices calculate indices based on prices collected by conducting price surveys.

Sampling

- Target outlets and target products are selected for month t and the relevant prices and quantities are sampled.
  The rule is described above.

- The quantity is the total volume of monthly sales of the given target product at the target outlet in month t.

- With regard to price,
  the daily unit price is determined based on daily sales records in scanner data,
  and the mode value for month t is used as the price for that month.

Satoshi Imai s2.imai@soumu.go.jp

Price count ratio and no of observations

Category	sampling A		sampling B		sampling C		sampling D	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
234001	0.995	0.015	0.998	0.008	0.998	0.009	0.995	0.010
107008	0.848	0.171	0.838	0.172	1.000	0.000	1.000	0.002
137001	1.000	0.000	0.995	0.015	0.999	0.004	0.999	0.004
031031	1.000	0.000	0.995	0.015	0.995	0.017	0.997	0.011
046001	0.995	0.012	0.997	0.009	1.000	0.000	1.000	0.000
041001	1.000	0.000	1.000	0.002	1.000	0.000	1.000	0.000
191001	1.000	0.004	1.000	0.002	1.000	0.000	1.000	0.000
212001	0.903	0.116	0.872	0.091	0.873	0.114	0.874	0.124
610001	0.999	0.005	0.998	0.005	0.996	0.008	0.998	0.004
612042	1.000	0.003	0.993	0.015	0.989	0.010	0.990	0.006
620001	1.000	0.000	0.998	0.004	0.985	0.009	0.978	0.015
818001	1.000	0.000	1.000	0.000	0.985	0.009	0.983	0.006
601021	1.000	0.000	1.000	0.000	0.986	0.009	0.990	0.006
Total No. of observations	118,343		235,574		118,988		237,706	

Satoshi Imai s2.imai@soumu.go.jp

Price Index Formulae (The elementary indices)

The period t price and quantity vectors are defined as

$p^t \equiv [p_1{}^t,\ldots,p_N{}^t]$ and $q^t \equiv [q_1{}^t,\ldots,q_N{}^t]$ respectively.

Carli:

$$P_C(p^0,p^t,q^0,q^t) \equiv \sum_{n=1}^{N} (1/N)(p_n{}^t/p_n{}^0)$$

Dutot:

$$P_D(p^0,p^t,q^0,q^t) \equiv \sum_{n=1}^{N} (1/N)(p_n{}^t)/ \sum_{n=1}^{N} (1/N)(p_n{}^0)$$

Jevons:

$$P_J(p^0,p^t,q^0,q^t) \equiv \prod_{n=1}^{N} (p_n{}^t/p_n{}^0)^{1/N}$$

Weighted Jevons:

$$P_J(p^0,p^t,q^0,q^t) \equiv \prod_{n=1}^{N} (p_n^t / p_n^0)^{s_n^0}$$

where $s_n{}^0 \equiv p_n{}^0 q_n{}^t / \sum_{j=1}^{N} p_j{}^0 q_j{}^t$

Satoshi Imai s2.imai@soumu.go.jp

Price Index Formulae (The elementary indices)

The period t price and quantity vectors are defined as

$p^t \equiv [p_1{}^t,...,p_N{}^t]$ and $q^t \equiv [q_1{}^t,...,q_N{}^t]$ respectively.

Törnqvist:

$$\ln P_T(p^0,p^t,q^0,q^t) \equiv \sum_{n=1}^N (1/2)(s_n{}^0+s_n{}^t)\ln (p_n{}^t/p_n{}^0)$$

Lowe:

$$P_{Lo}(p^0,p^1,q) \equiv \sum_{n=1}^N (p_n{}^t/p_n{}^0)\, s_n{}^{0b}$$

where $s_n{}^{0b} \equiv p_n{}^0 q_n{}^b / \sum_{j=1}^N p_j{}^0 q_j{}^b$

The weight reference period for the Lowe index is one year prior to the price reference period.

Satoshi Imai s2.imai@soumu.go.jp

Price Index Formulae (The elementary indices)

The period t price and quantity vectors are defined as

$p^t \equiv [p_1^t,...,p_N^t]$ and $q^t \equiv [q_1^t,...,q_N^t]$ respectively.

Laspeyres:

$$P_L(p^0,p^t,q^0) \equiv \sum_{n=1}^N p_n^t q_n^0 \,/\, \sum_{n=1}^N p_n^0 q_n^0$$

Paasche:

$$P_P(p^0,p^t,q^t) \equiv \sum_{n=1}^N p_n^t q_n^t \,/\, \sum_{n=1}^N p_n^0 q_n^t$$

Fisher:

$$P_F^t \equiv [P_L^t\ P_P^t]^{1/2}$$

Satoshi Imai s2.imai@soumu.go.jp

Index Calculation

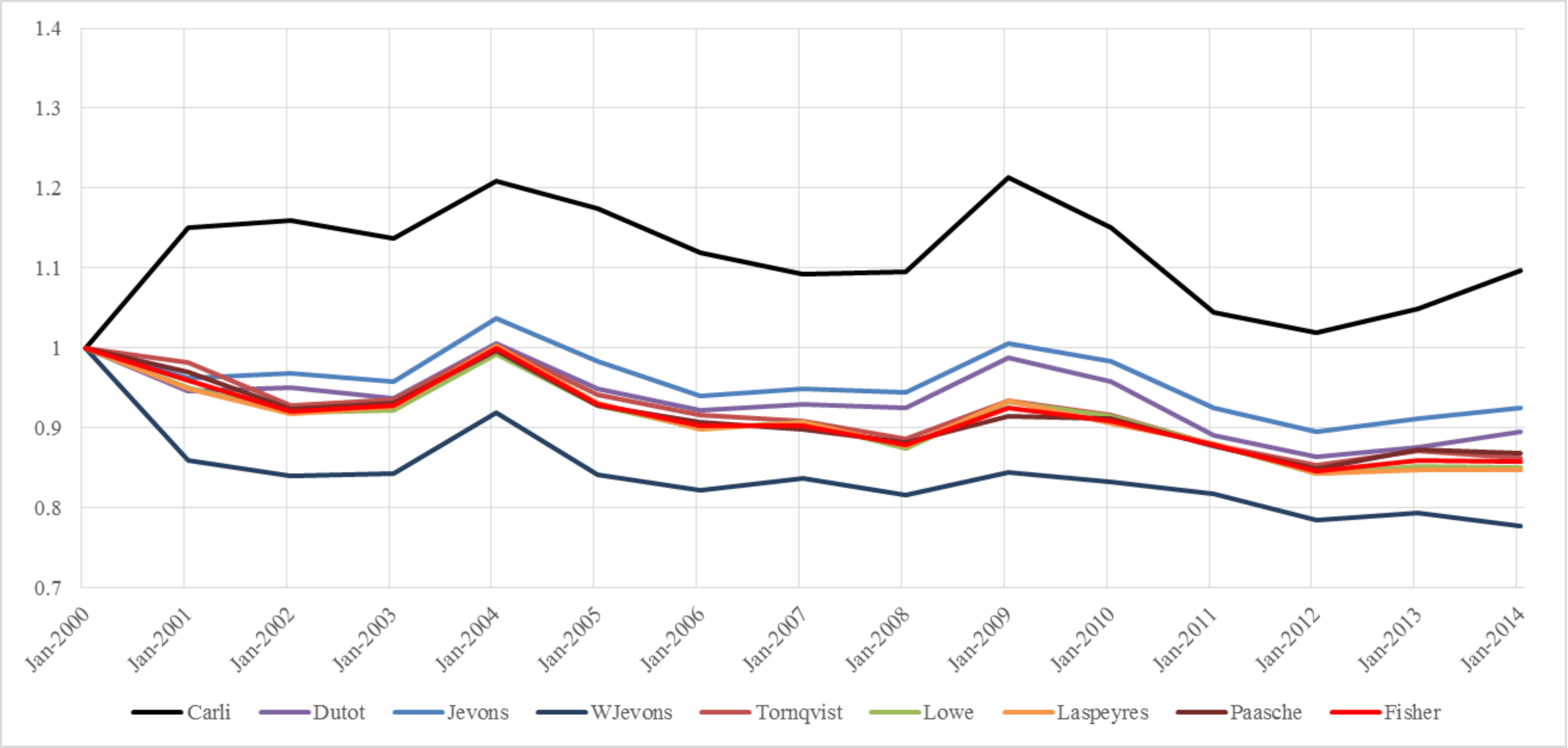
- Taking January 2000 as the reference period and using direct comparison without performing quality adjustment, we calculated year-over-year indices from January 2000 to January 2014.

- We used the Lowe (annual weight) and Laspeyres formulae to aggregate item level indices into national indices.

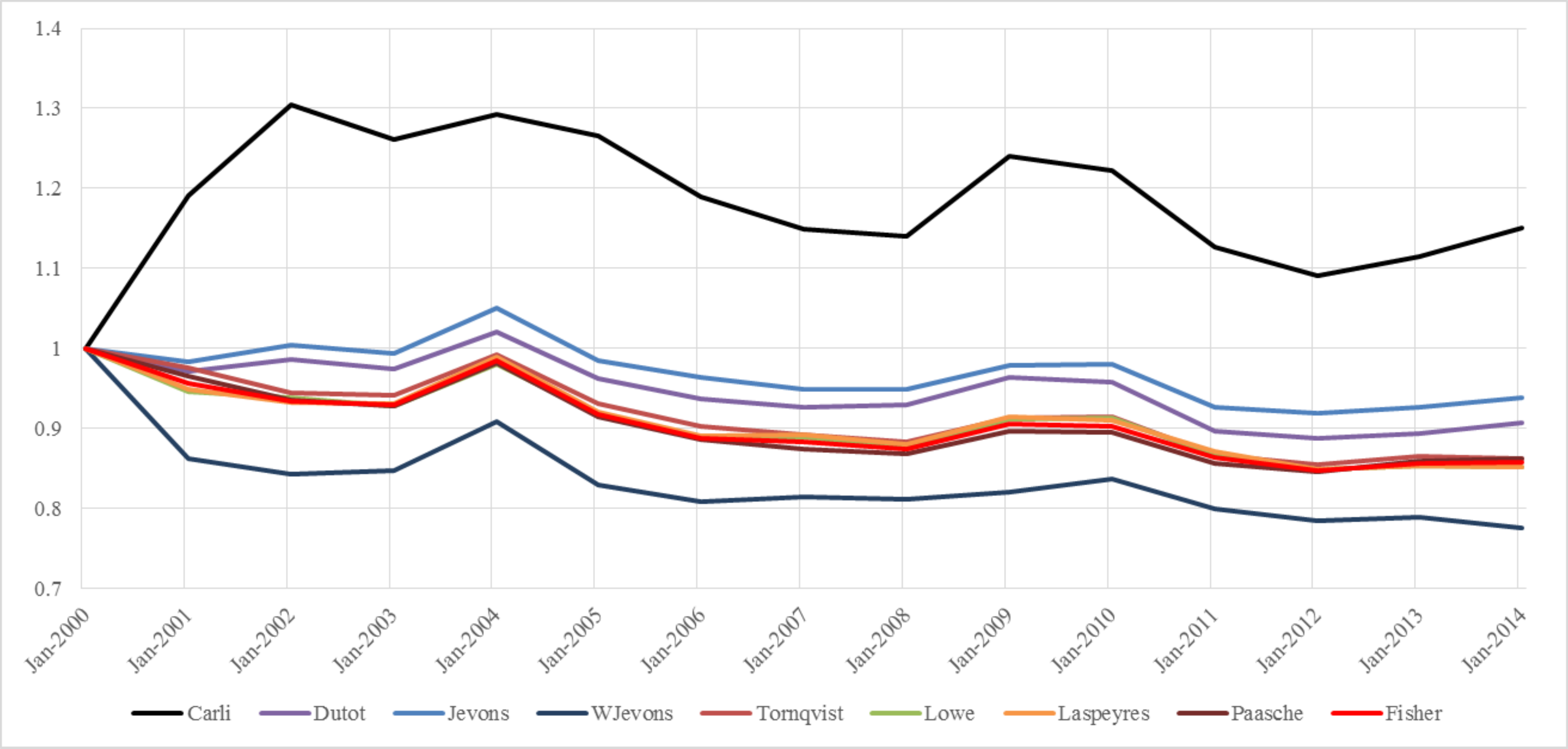
- In order to conduct comparison at the aggregation stage, we calculated a two-stage Fisher index, too.

Satoshi Imai s2.imai@soumu.go.jp

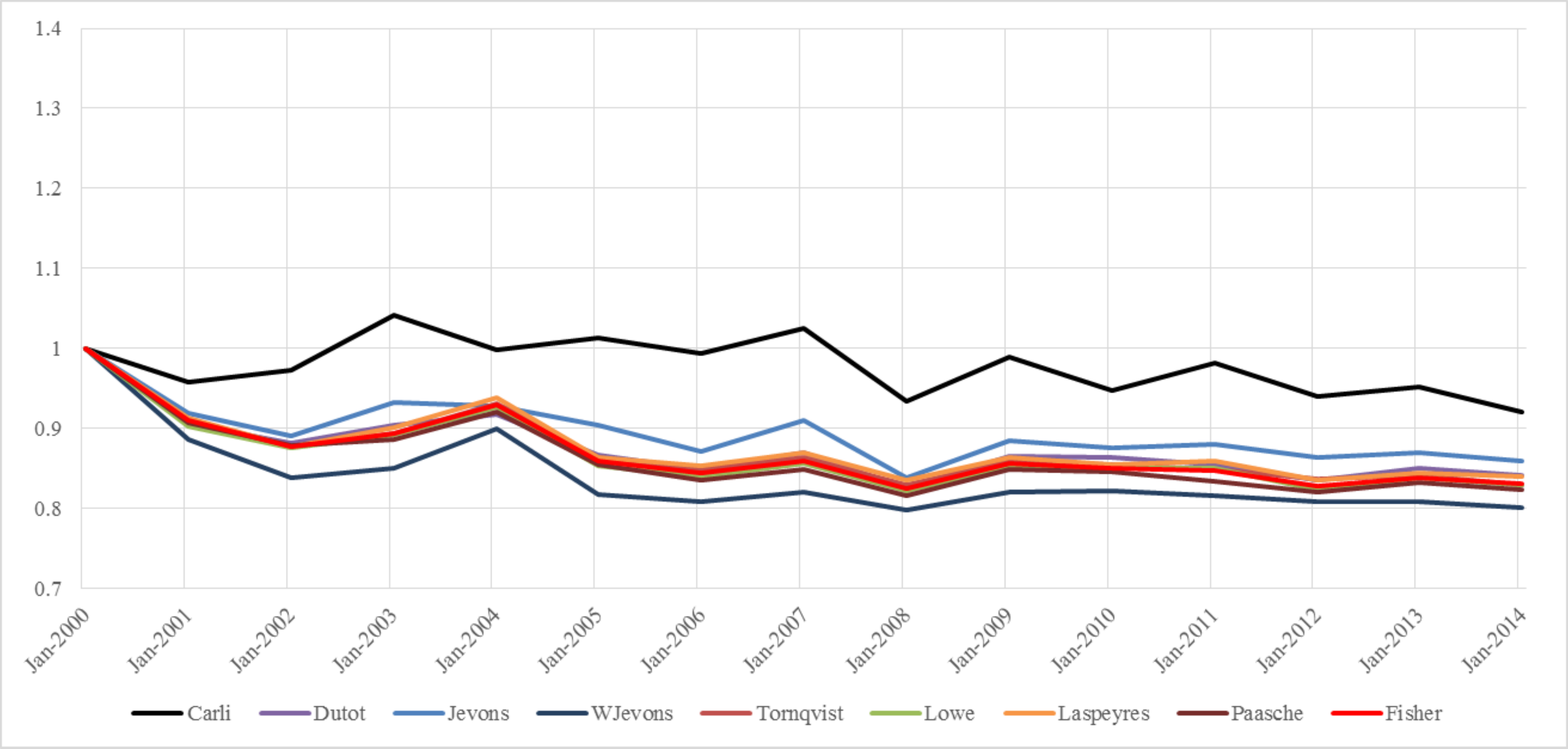
Laspeyres Aggregation Year-over-Year Index —Top 5 Items, 10 Outlets



Laspeyres Aggregation Year-over-Year Index —Top 5 Items, 20 Outlets

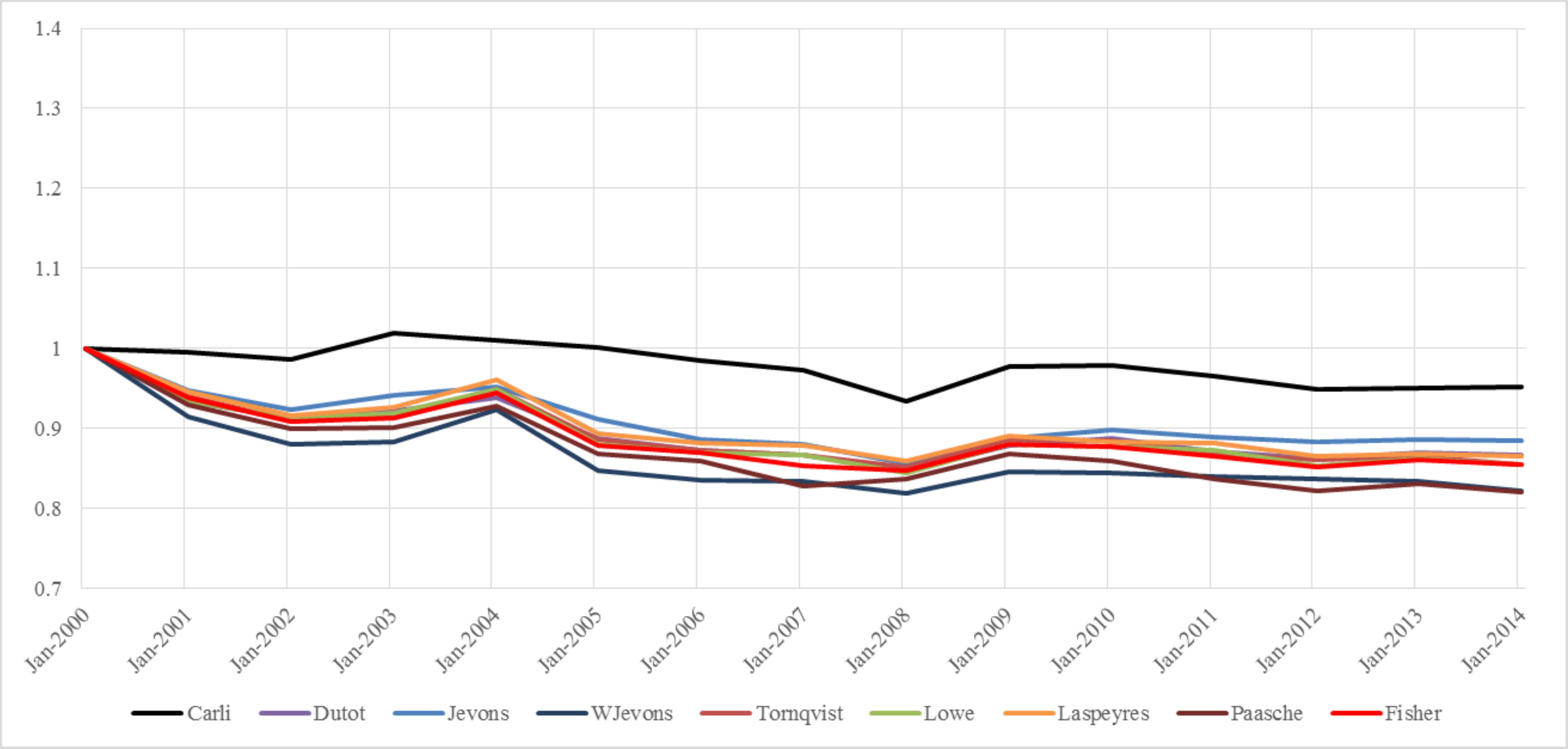


Laspeyres Aggregation Year-over-Year Index —Top 1 Items, 50 Outlets

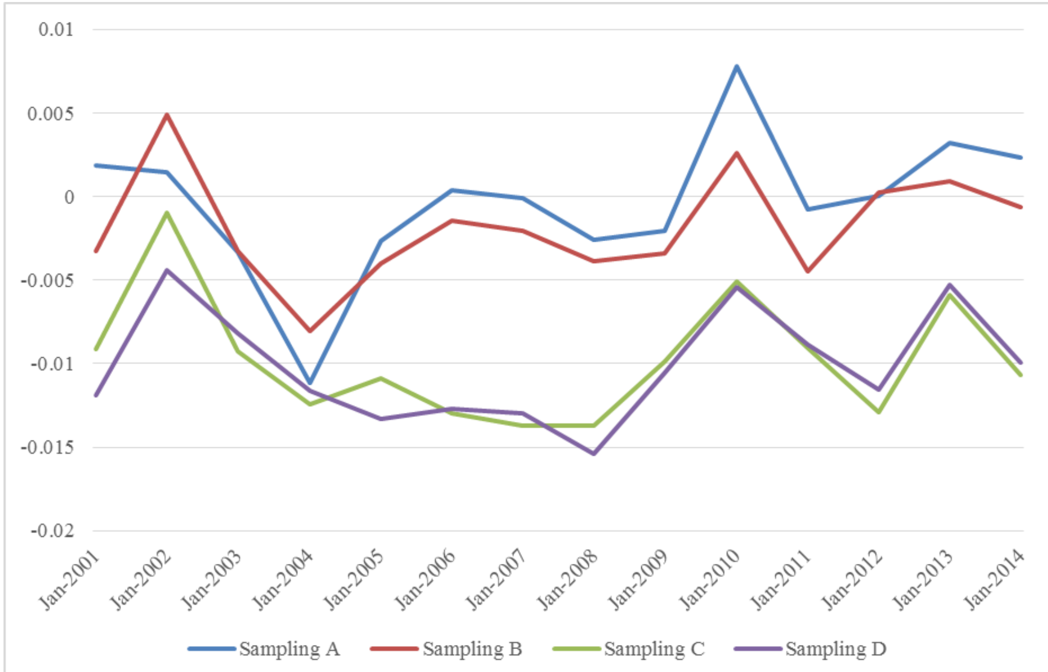


Satoshi Imai s2.imai@soumu.go.jp

Laspeyres Aggregation Year-over-Year Index —Top 1 Items, 100 Outlets



The Difference between Lowe and Laspeyres Price Indices



Sampling C and D, the difference has a negative value, which rejects the idea that it is 0.

This is an unusual result, given that the value would normally be expected to be positive. It may be due to the fact there was deflation in Japan during the sample period covered by this paper, so price trends differed from the normal assumptions regarding long-term price trends.

Satoshi Imai s2.imai@soumu.go.jp

Biases Based on Choice of Elementary Index

- Elementary index biases are defined as the differences between calculation results using the Fisher price index and the other elementary indices.
- Since the elementary indices are calculated using the item level indices before performing upper-level aggregation, the biases were calculated for the 13 items for month t.
- We calculated the average and standard deviation for these biases by item and created 14 series from January 2001 to January 2014. Furthermore, we calculated the average and standard deviation for the 14 time series.

- In this manner, we tested what kind of differences occur in each elementary index's divergence from the Fisher index as a result of changing the sampling method.

Satoshi Imai s2.imai@soumu.go.jp

Item Level Sampling Method Bias

- The sampling method bias is defined as the difference obtained by subtracting the Sampling D (top product at 100 outlets) Fisher index from the Fisher index for each of the other sampling methods (A, B, and C).
- We calculated the average and standard deviation of the difference for each item for month t.
- We also created 14 series from January 2001 to January 2014 and calculated the average and standard deviation for them.

- In this manner, we tested the structure of sampling method-based item level biases by evaluating differences between sampling methods for Fisher index item levels.

Aggregation-Level Sampling Method Bias

- Next, we compared sampling method biases at the aggregation level.
- We compared aggregate indices calculated using the two-stage Fisher method.
- We compared differences between indices calculated with each sampling method, taking Sampling D (top product at 100 outlets) as the reference.
- We created 14 series from January 2001 to January 2014 and calculated the averages and standard deviations.
- In this manner, we evaluated aggregation-level sampling method biases and tested whether these biases could be canceled out at the aggregate level.

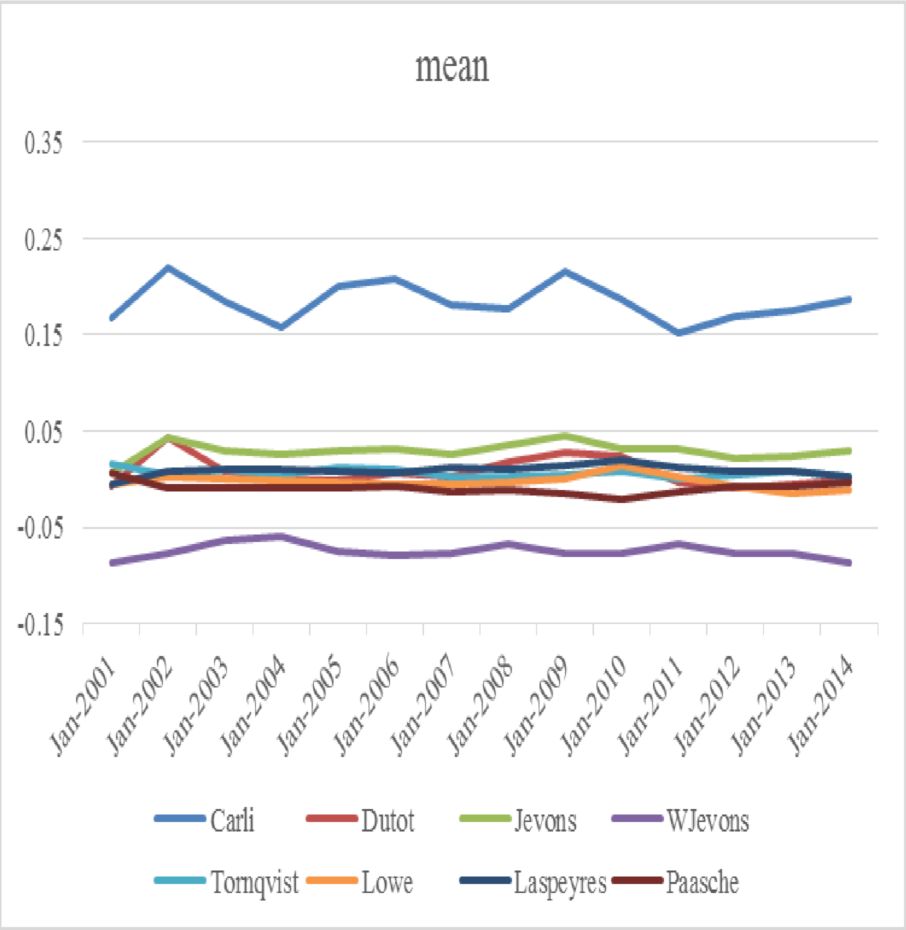
Satoshi Imai s2.imai@soumu.go.jp

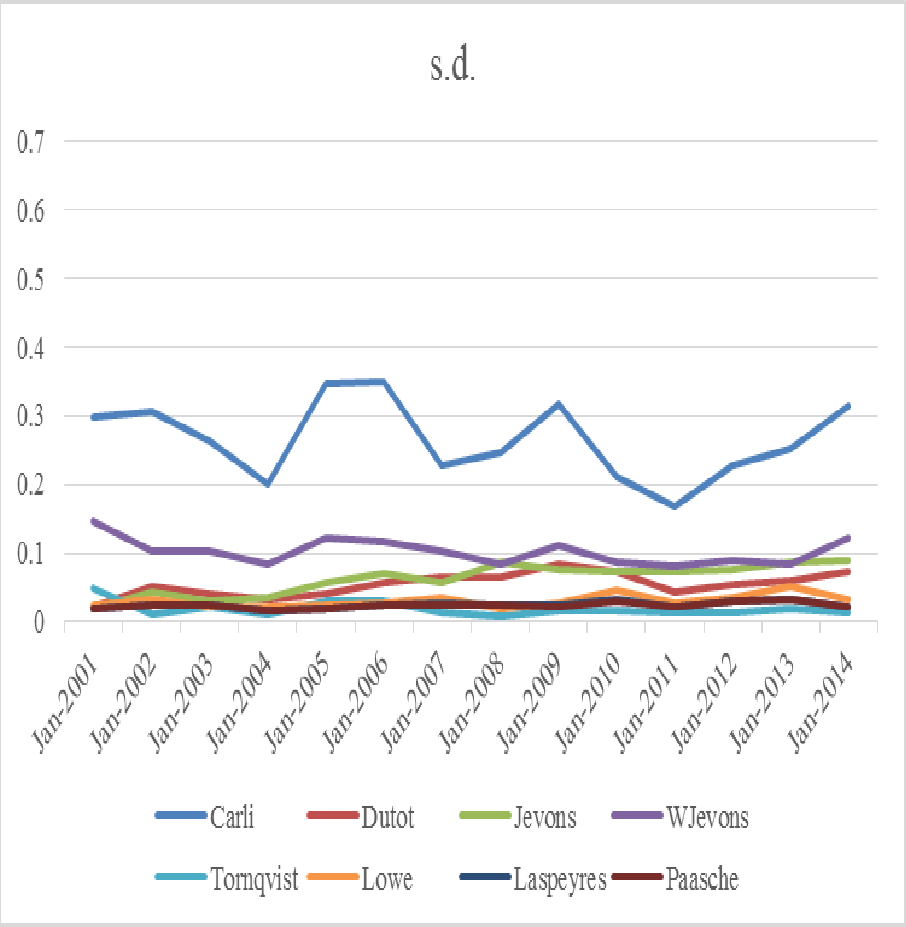
Evaluation of Aggregation-Level Sampling Method Bias

- Finally, in order to evaluate whether the magnitude of sampling method biases at the aggregation level is at a level that can be ignored.
- Upper-level substitution bias may be defined as the difference between the upper-level Laspeyres index calculated using Fisher as the elementary index and the index calculated using the two-stage Fisher method.
- We also calculated the difference between the aggregate Lowe index calculated using Fisher as the elementary index and the index calculated using the two-stage Fisher method.
- In this way, we tested whether aggregation-level sampling method biases were sufficiently less than upper-level substitution biases.

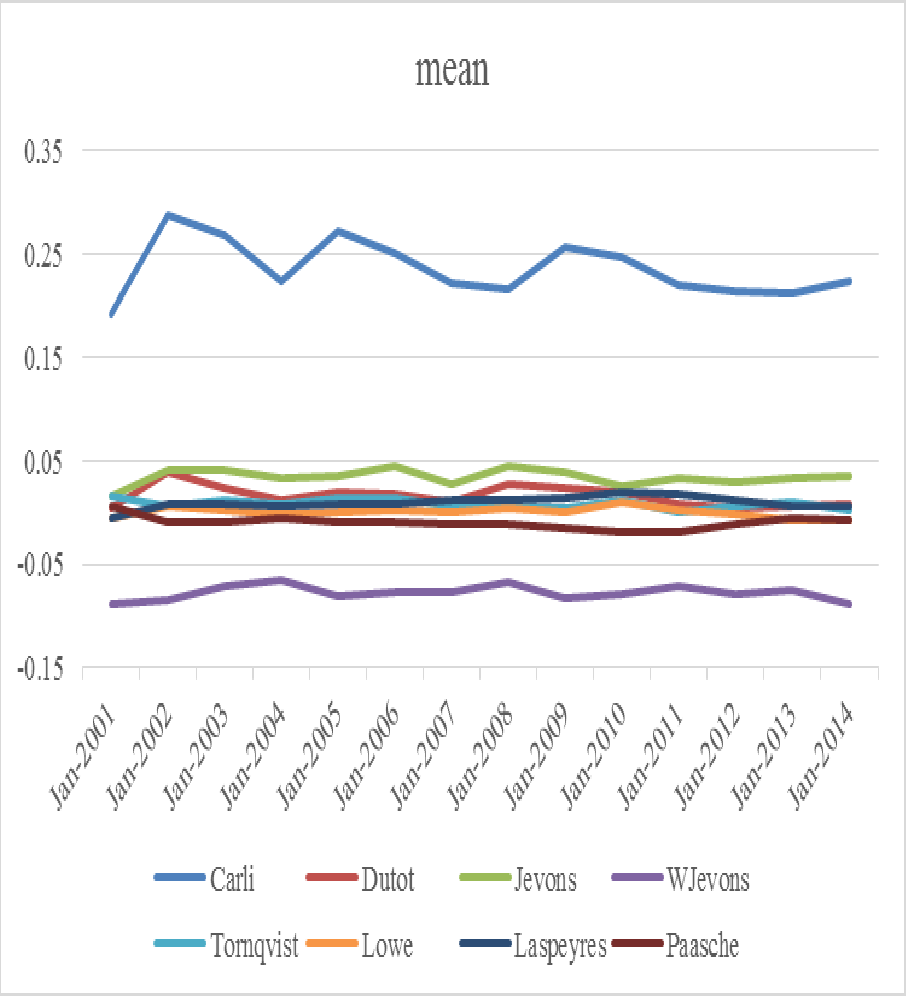
Satoshi Imai s2.imai@soumu.go.jp

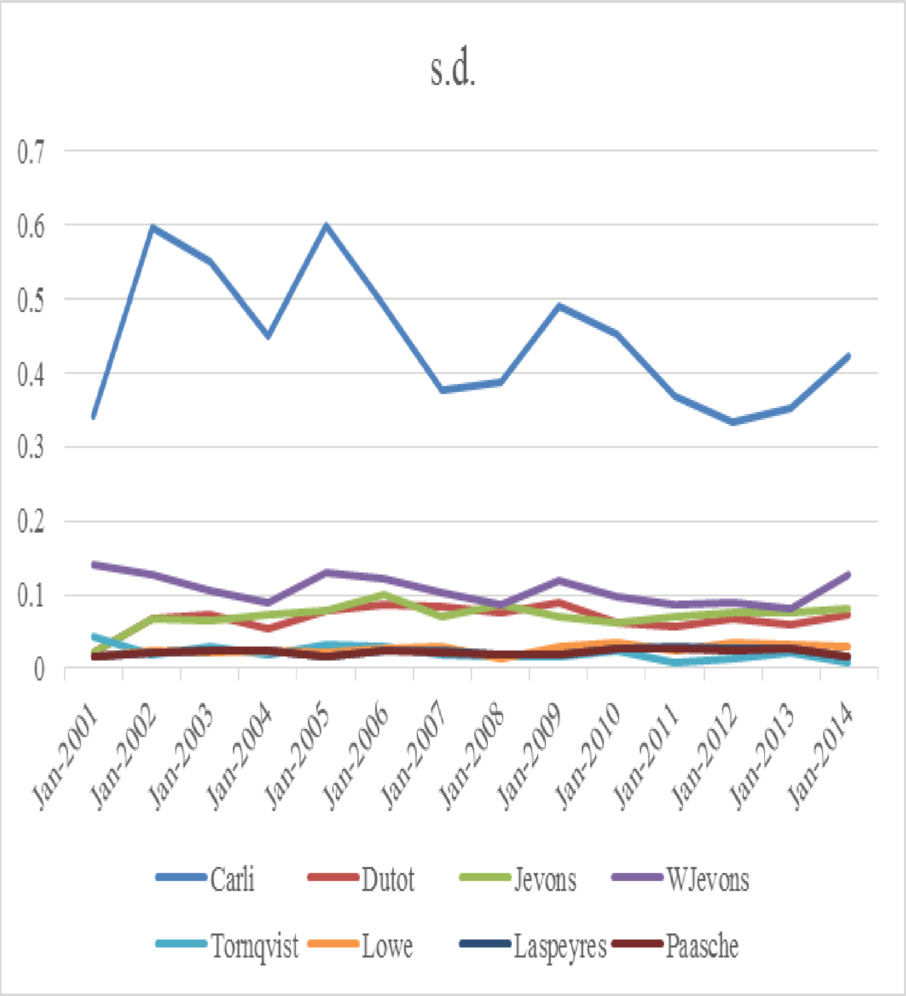
Elementary Index Bias by Formula over Time (Mean, S.D.)—Top 5 Items, 10 Outlets



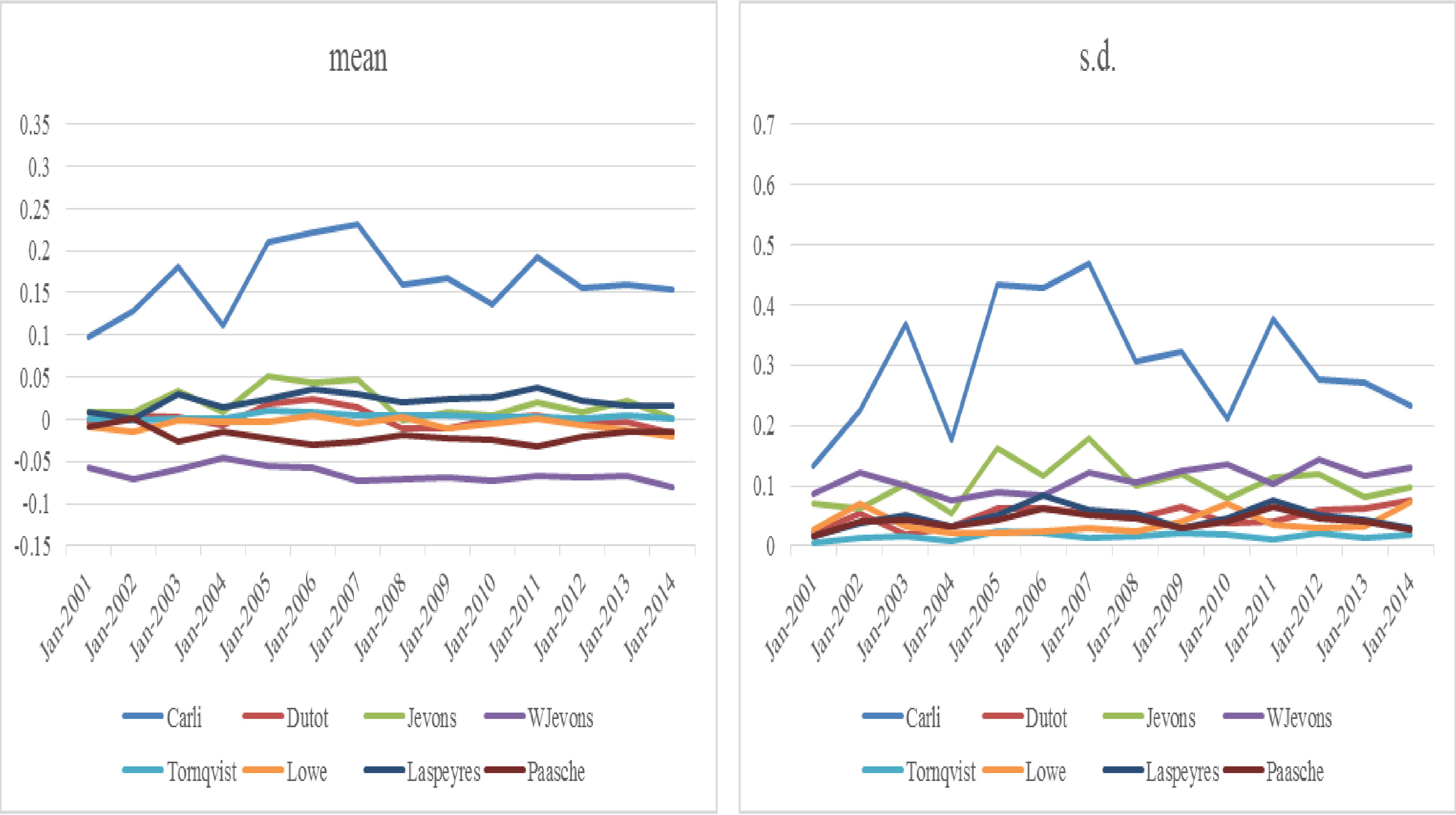


Elementary Index Bias by Formula over Time (Mean, S.D.)—Top 5 Items, 20 Outlets

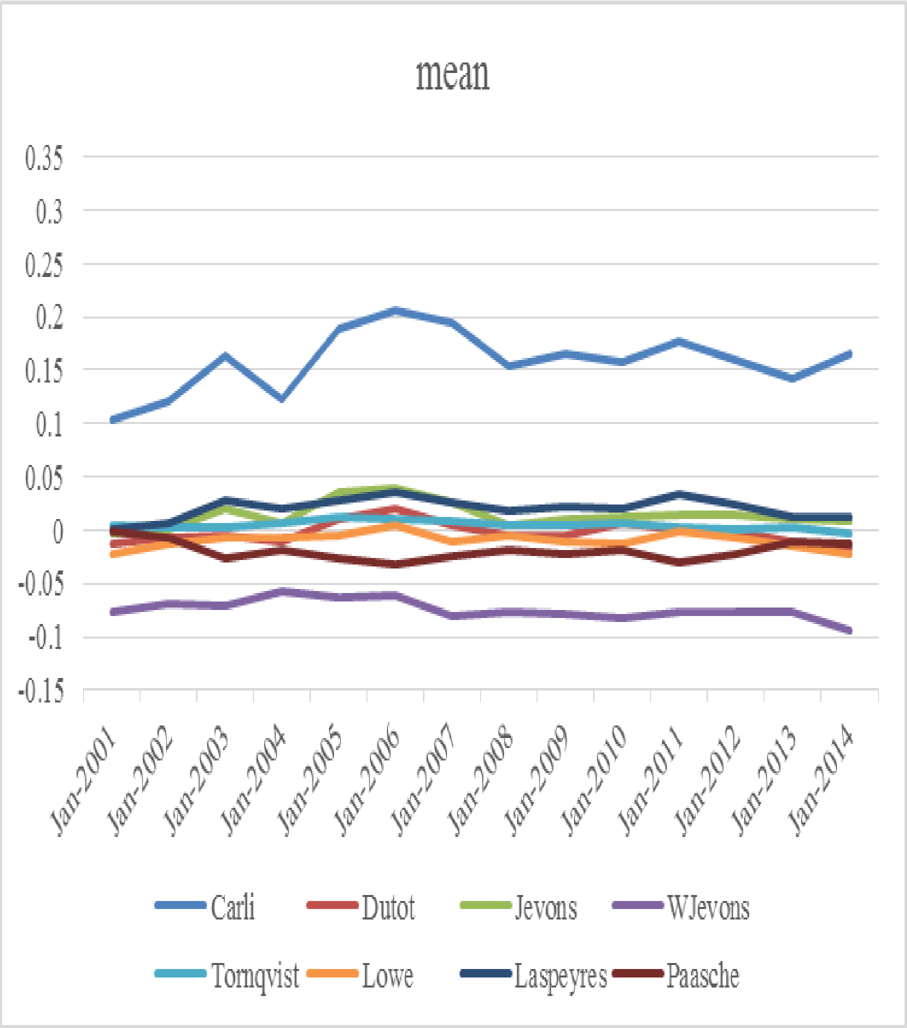


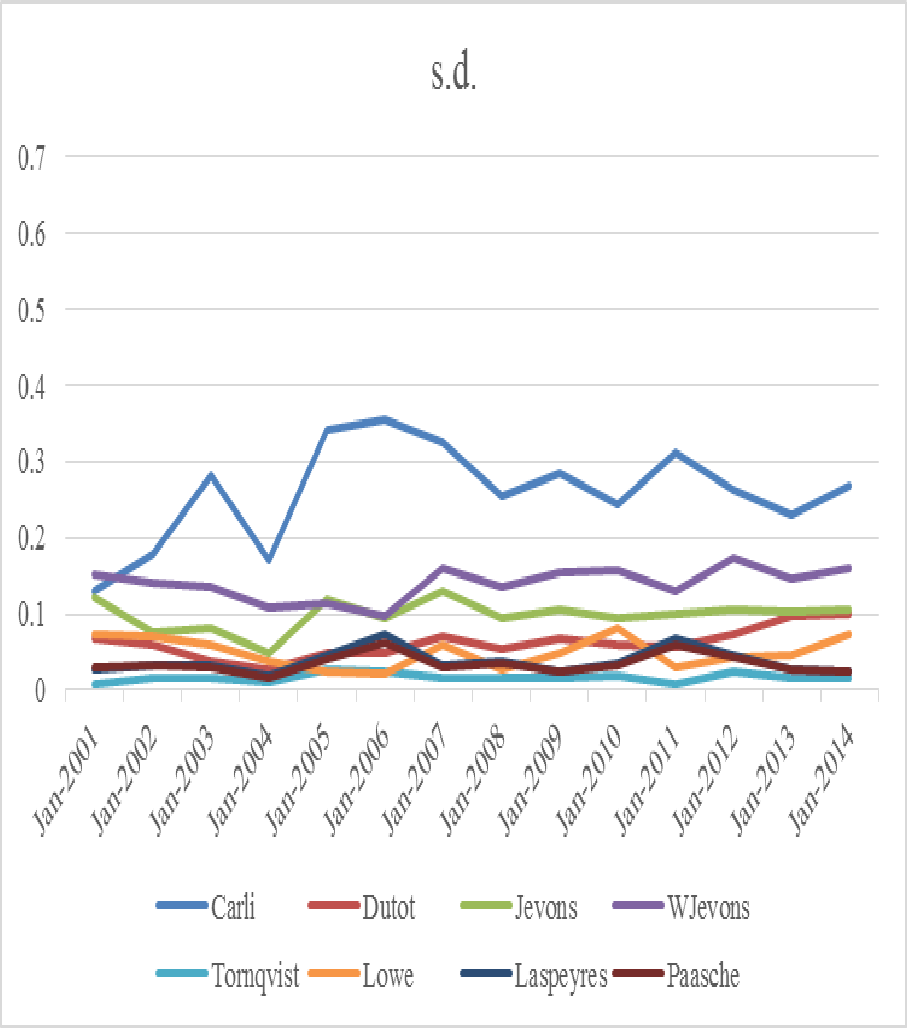


Elementary Index Bias by Formula over Time (Mean, S.D.)—Top 1 Item, 50 Outlets



Elementary Index Bias by Formula over Time (Mean, S.D.)—Top 1 Item, 100 Outlets





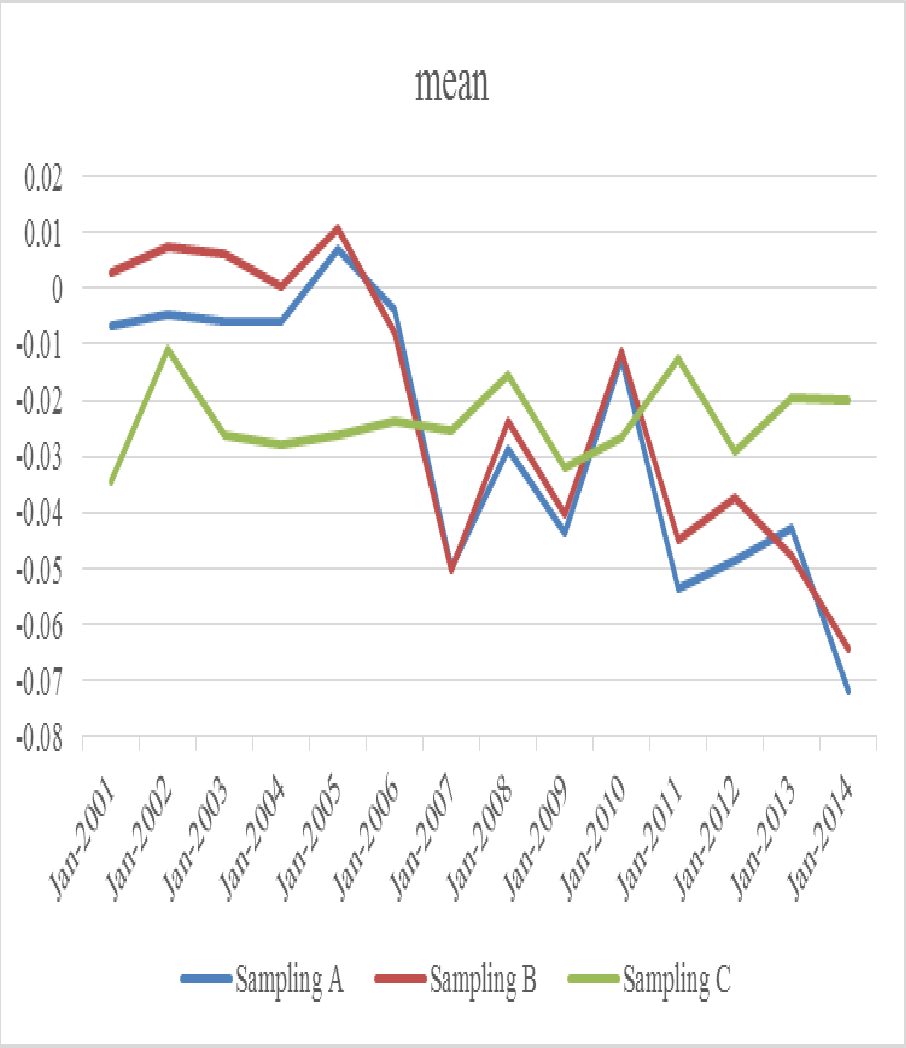
Biases Based on Choice of Elementary Index

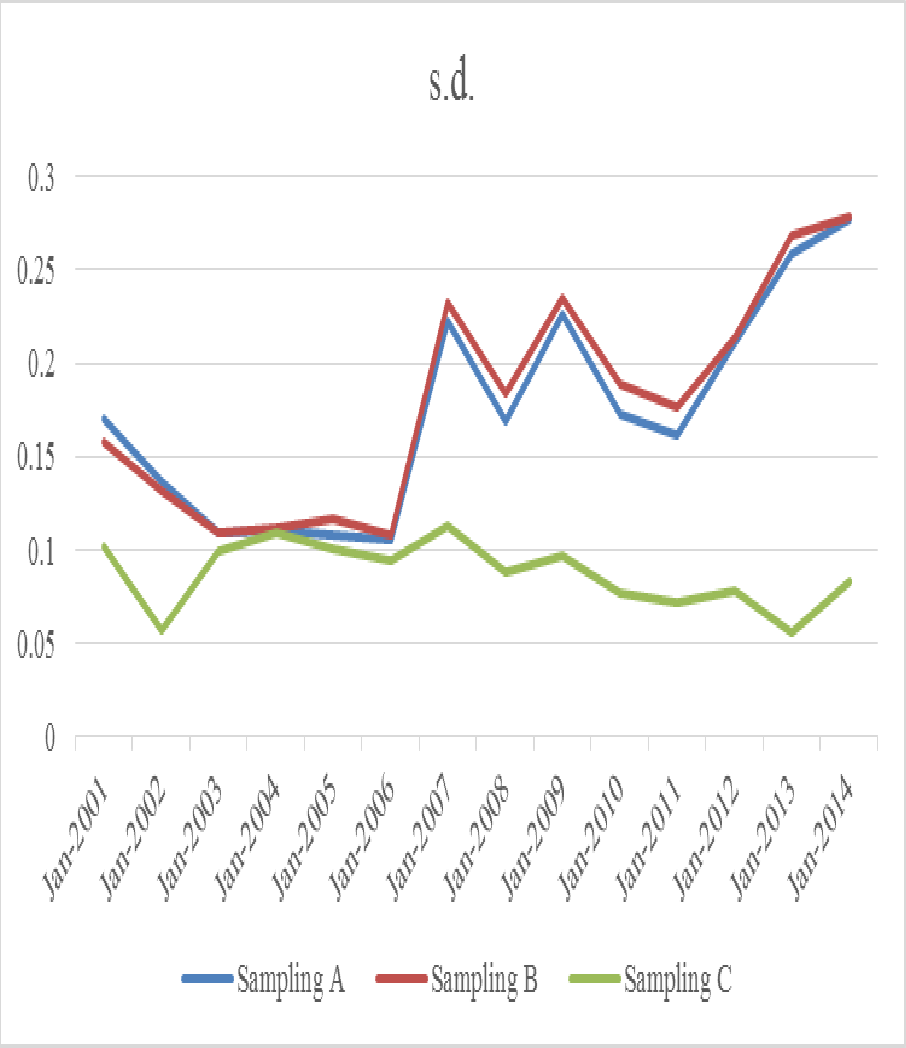
- The sampling method produces biases independent of the choice of elementary index.
- In cases where only the top product is sampled, if the number of observations is increased, both the mean difference from the Fisher index and the standard deviation become smaller.
- But with the method of sampling the top five products, if the number of observations is increased, both the mean difference and the standard deviation become larger.

- In particular, the difference between selecting the top product only and selecting the top five products seems to have a greater effect than changing the sample size.

Satoshi Imai s2.imai@soumu.go.jp

Sampling Method Bias over Time (Mean, S.D.)

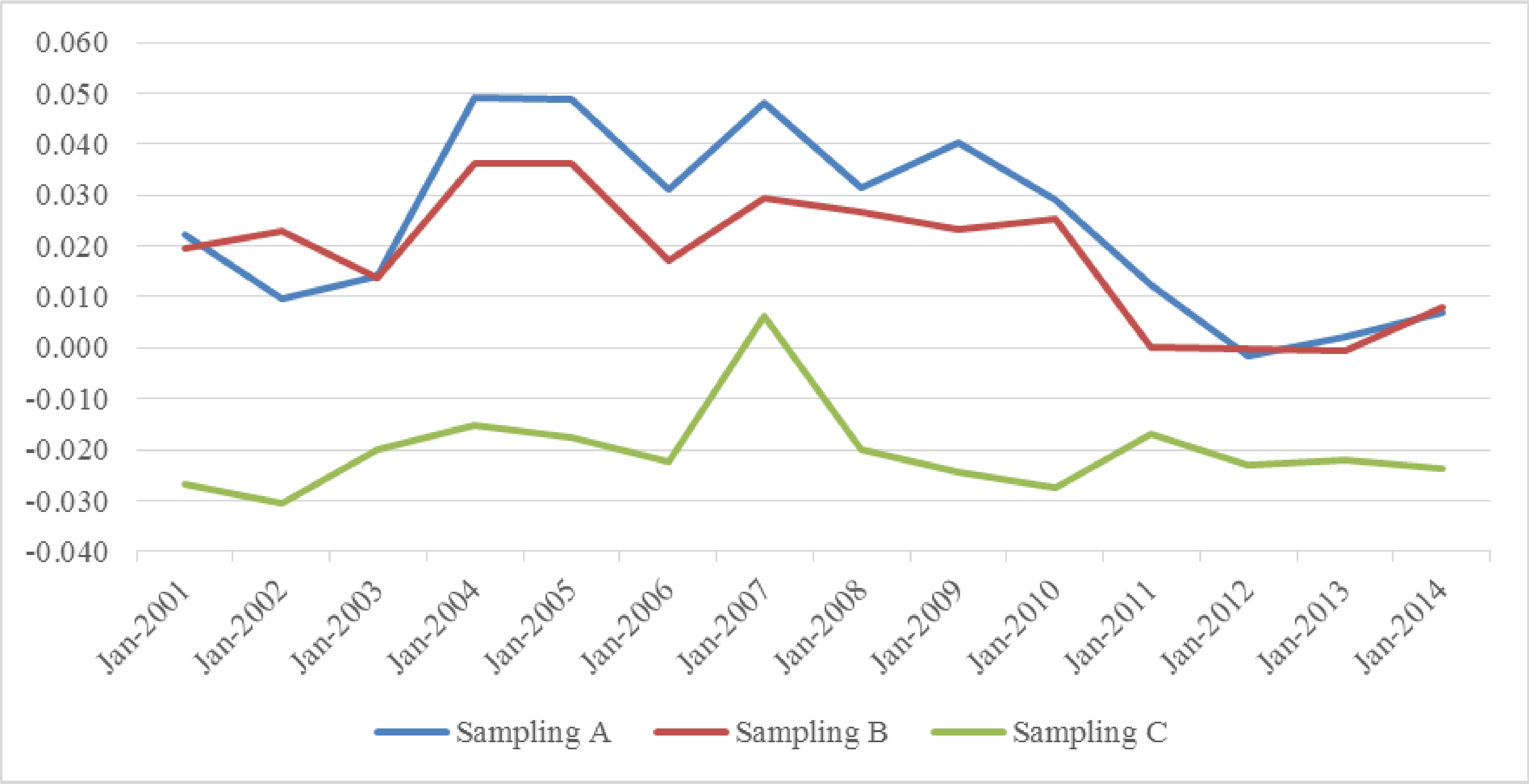




Item Level Sampling Method Bias

- When sampling only the top product, increasing the number of samples means broadening the scope to include more outlets with relatively low popularity.
- As a result, in Sampling C, which uses only the 50 top outlets, there is a relatively higher proportion of large-scale stores that adopt a low-margin, high-volume price strategy, which has a downward effect compared to the method which samples 100 outlets.
- with the method of sampling the top five products, we may consider that the variation in products has a significant effect on the results.
- This demonstrates that the selection of the number of representative products is important.

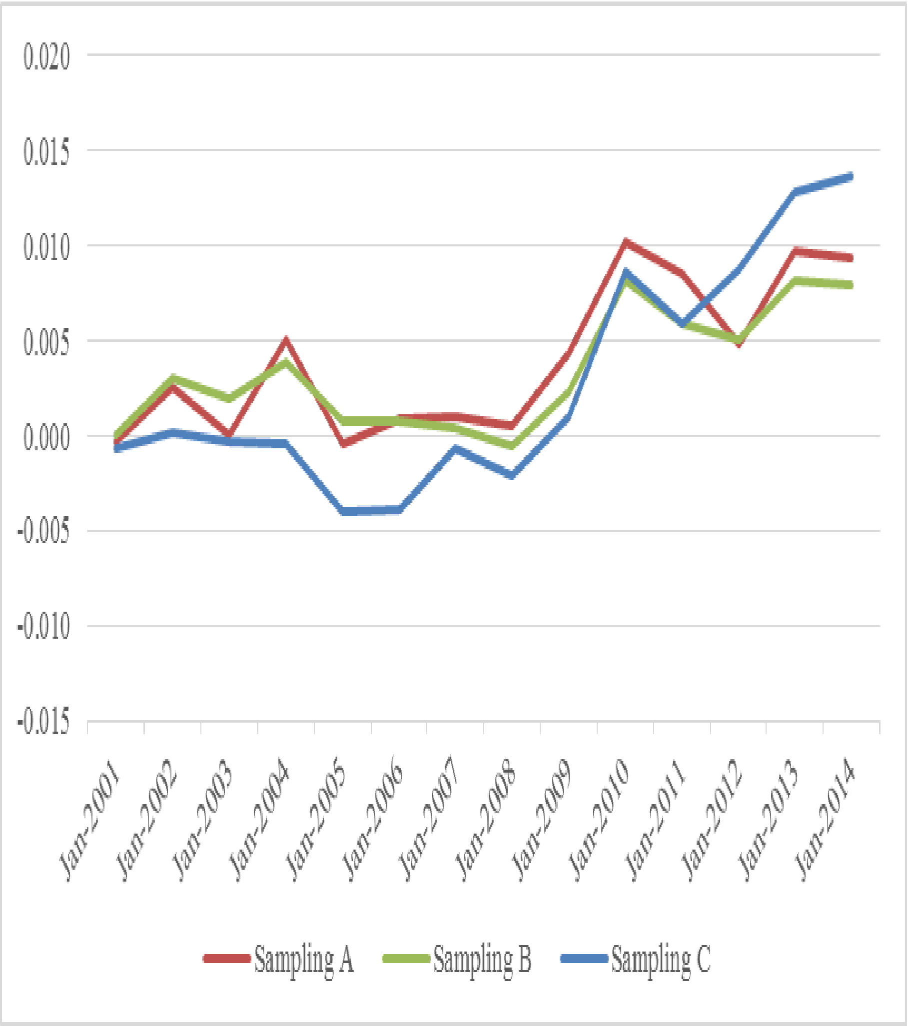
Sampling Method Bias over Time (Aggregation Level)

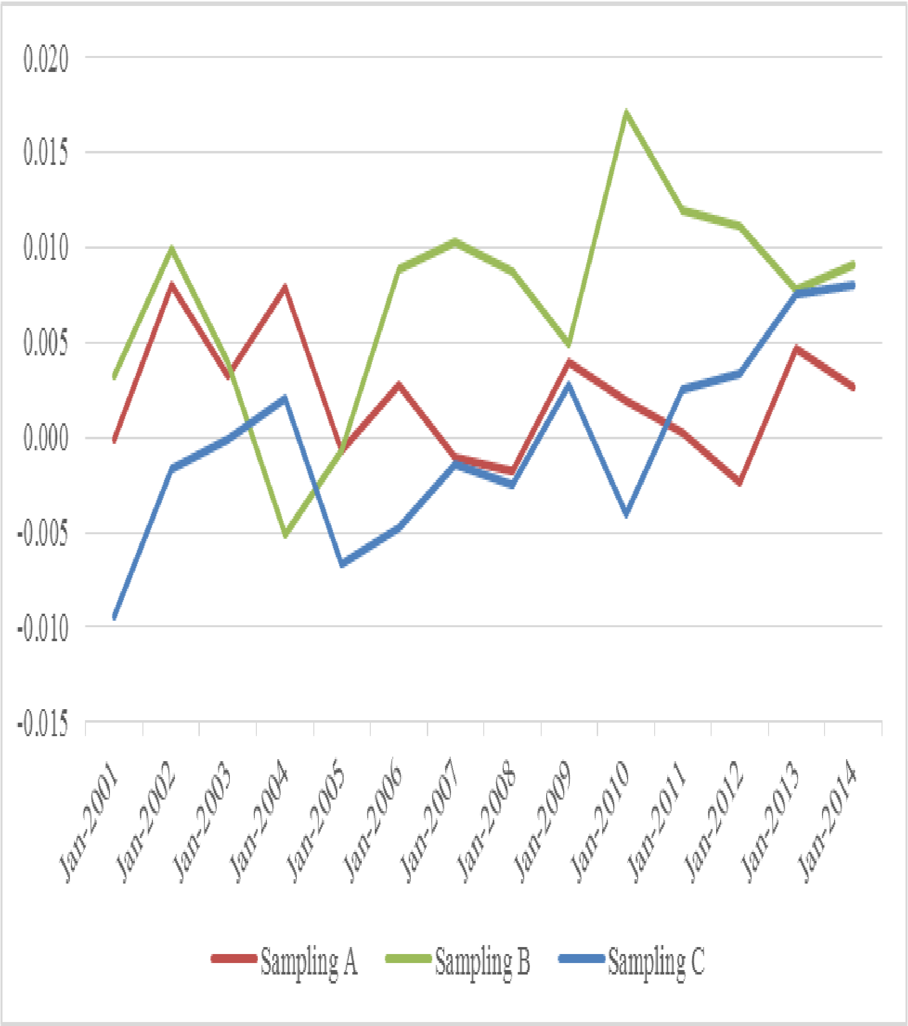


Aggregation-Level Sampling Method Bias

- The sampling method bias is not canceled out between products and remains even in upper-level indices
  But the cause does not seem to be fixed.
- For sampling the top product only,
  if the sample size is small, prices at high-popularity outlets which may adopt a low-margin, high-volume pricing strategy have a greater weight, as a result of which a downward bias is generated.
- For sampling the top five products,
  the bias is not stable at either the item level or the aggregate level.
  This may be because the variation in the sampled products affects the results.

Upper-Level Substitution Bias over Time





Satoshi Imai s2.imai@soumu.go.jp

Evaluation of Aggregation-Level Sampling Method Bias

- Regardless of the sampling method or elementary index, upper-level substitution bias mostly falls within the range of -0.01 to 0.01
- And the magnitude is large in relation to the average differences seen in the aggregation level sampling method bias (0.025 for Sampling A, 0.018 for Sampling B, and -0.020 for Sampling C)

- So even at the aggregation level, sampling method-based biases are at a level that cannot be ignored.

Conclusions:

- Sampling methodologies for elementary indices that correspond to more lower-level prices may have effects that cannot be ignored at both the item level and at higher levels of aggregation.

- In the case of sampling the top product only, there is a possibility of underestimation if there are few outlets. if the sample size is small with this method, a downward bias may be generated due to prices at high-popularity outlets.
- There is a possibility of over-estimation if one samples the top five products from a few upper-level outlets, but we saw that with methods that involve sampling the top five products, biases do not stabilize over time.

Conclusions:

- We propose that in order to discuss lower-level biases, it is necessary to clearly separate discussion based on the choice of elementary index from discussion relating to the sampling method.
- Large-scale sources such as scanner data are recognized for their value in terms of the breadth of their coverage, but when considered from the standpoint of representativeness, price information for products with a low level of representativeness will consistently be included.
- This does not mean that scanner data cannot be used for price indices. Rather, it suggests that it is essential to fully understand the attributes of different types of data when using them.