Compilation of Experimental Price Indexes using Big Data and Machine Learning

— A Comparative Analysis and Validity Verification —

Meeting of the Group of Experts on Consumer Price Indices
held in Palais des Nations, Genève

May 7-9, 2018

Kimiaki Shinozaki

Bank of Japan

1. Introduction

1. Introduction

- A price index is constructed with a primary aim of understanding fluctuations in general price levels by indexing the *constant-quality* price of targeted representative products with the price at the base point in time as 100.
- In order to ensure representativeness of surveyed products, it is necessary to perform a **change of sample prices** at an appropriate frequency, and adopting strong-seller goods to be surveyed.

Question

How should differences between old and new products that do not have identical quality specifications be processed?

1. Introduction

- Price statisticians have adopted the method of processing quality differences between old and new products by splitting the price differences as of the same point in time into **"price change due to quality changes"** and **"pure price change"**. By eliminating the former, the price index only reflects the latter.

$$\boxed{\text{Price difference between old and new product}} = \underbrace{\boxed{\text{Price change resulting from quality changes}}}_{\text{Excluded using quality adjustment}} + \underbrace{\boxed{\text{Pure price change}}}_{\text{Reflected into price index}}$$

- In recent years, accompanying advance of big data analysis, price statisticians and macroeconomists are challenging to compile price indexes using **scanner data** or **webscraped data**.

1. Introduction

- In this paper, we combine the following approaches:

***Traditional* approach**: Price index which is created by carrying out changes of sample prices reflecting the product life cycles and quality adjustments between old and new products.

and

***Non-traditional* approach**: Price index which is compiled by making use of big data and computing capabilities.

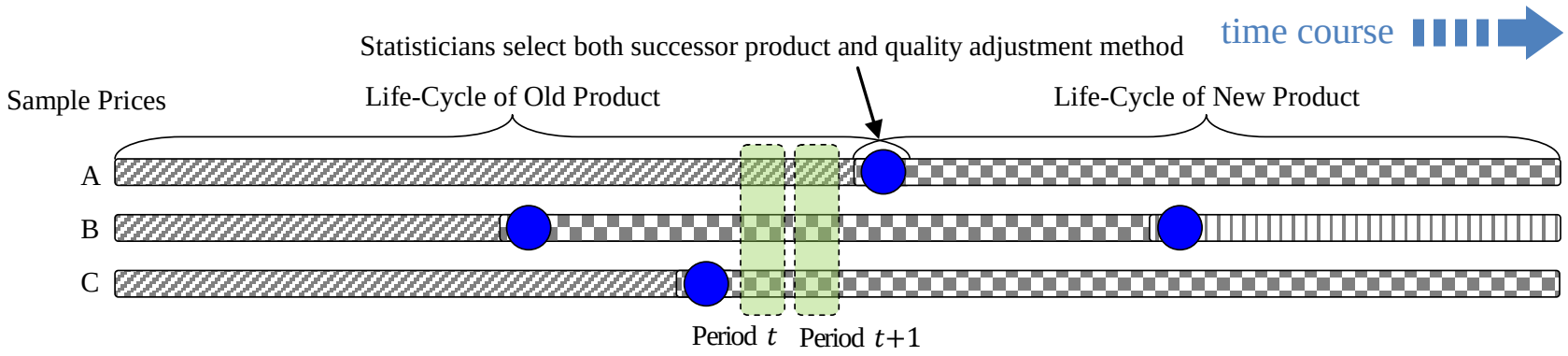
- We use **big data** obtained from Japan's leading price comparison website *Kakaku.com* and **machine learning methods** to imitate the know-how of price statisticians.

2. Comparison of Approaches for Compiling Price Indexes

2. Comparison of Approaches for Compiling Price Indexes

2-1. *Traditional* Approach of Price Statistics Agencies

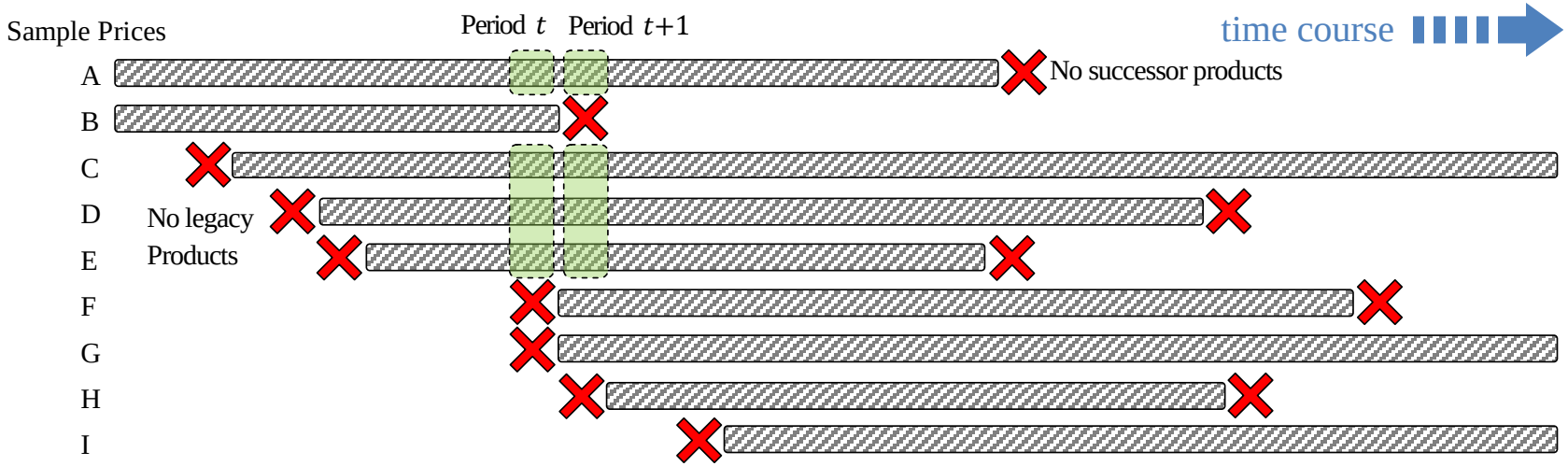
- Price statisticians select representative products to be surveyed, considering product specifications and data availability. At time of changing sample prices, they apply optimal quality adjustment to remove price change arising from changes of quality.
- Due to resources constraints at price statistics agencies and burden on reporting firms, the number of sample prices tends to be limited.



2. Comparison of Approaches for Compiling Price Indexes

2-2. *Non-Traditional* Approach Using Big Data

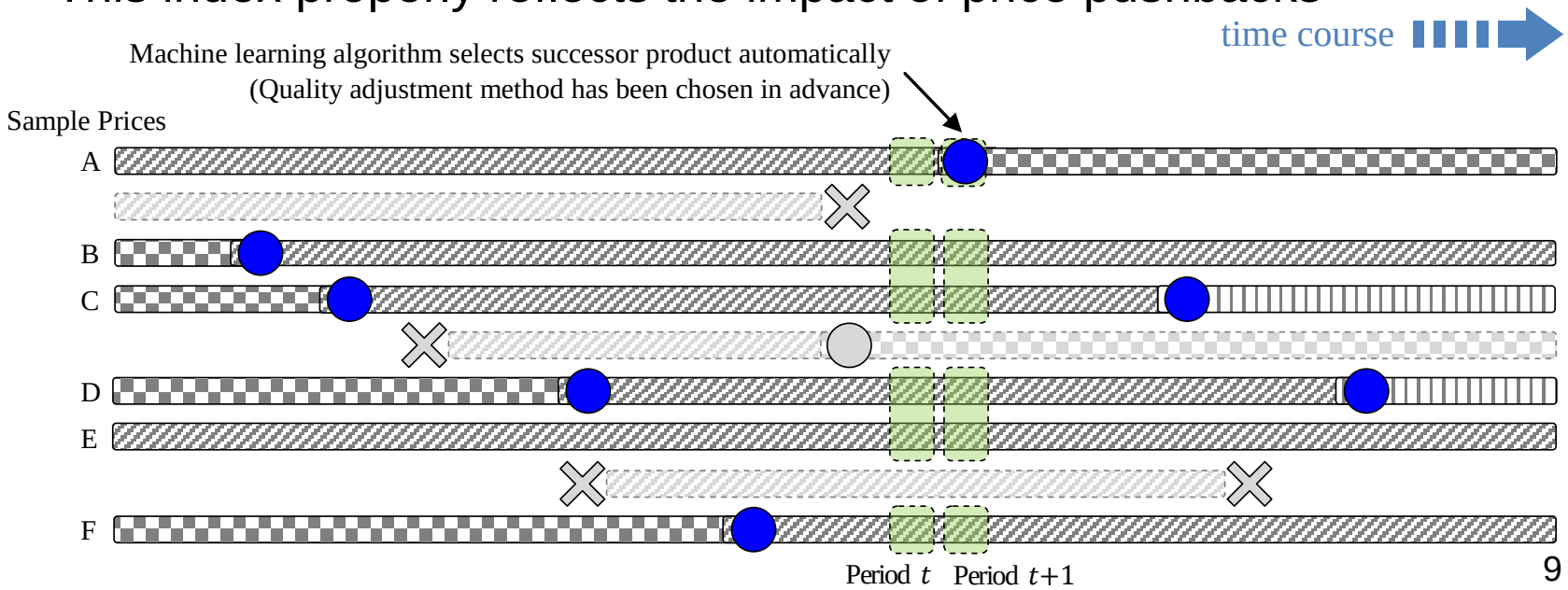
- Compiling price index by using big data and **Matched-Model Method** (MMM) which calculates the percentage change of price for products which exist in both *survey period* and *following period*.
- If price pushbacks are constantly conducted when launching new products, the index cannot properly reflect the impact of such price pushbacks, and may cause a downward bias.



2. Comparison of Approaches for Compiling Price Indexes

2-3. Approach We Take in This Paper

- In this paper, we combine *traditional* and *non-traditional* approaches.
- We developed a **supervised machine learning** algorithm which pairs legacy and successor products with high precision to conduct appropriate quality adjustment for given big data.
- This index properly reflects the impact of price pushbacks



2. Comparison of Approaches for Compiling Price Indexes

2-4. Significance of This Analysis

- Based on the empirical results presented at the last CPI Experts Group Meeting (Abe et al. (2016)), the Bank of Japan has introduced the **Webscraped Prices Comparison Method** (WSM) as one of the quality adjustment methods in Japan's PPI since 2017.
- WSM is used for home electronics which frequently conduct model changes accompanied by quality improvements, assuming 50% of webscraped retail price differences of old and new products are deemed to be quality improvement.
- In this paper, we verify the appropriateness of WSM by comparing with indexes compiled by **Hedonic Regression Method** (HRM), using the same data set.

Abe, N., Y. Ito, K. Munakata, S. Ohyama, and K. Shinozaki (2016), "Pricing Patterns over Product Life-Cycle and Quality Growth at Product Turnover: Empirical Evidence from Japan" Bank of Japan Working Paper Series No.16-E-5.

3. Making Old and New Product Pairs Using Machine Learning Method

3. Making Old and New Product Pairs Using Machine Learning Method

3-1. Outline of Data set

- Develop an unbalanced panel data sets by integrating the following:
  1. Product specifications: registered at the ***Kakaku.com*** between December 2012 and December 2015.
  2. Weekly average prices: registered at the paid ***Kakaku.com Trend Search*** between December 2013 and December 2015.
- Coverage:

  Home electrical appliances: 8 commodities

  Digital consumer electronics: 12 commodities
- Data Volume:

  Number of products: 4,500

  Size of panel data: 150,000

  Total data volume: 5.6 million

3. Making Old and New Product Pairs Using Machine Learning Method

3-2. Creation of Product Pairs

- We assemblage product pair data by generating combinations of two products from the data set. Then we narrow down the pairs by imposing the following three necessary conditions:

Necessary conditions to compose old and new product pairs

1. The release date of the new product is after the old one.
2. The old and new products are made by the same manufacturer.
3. Sales interval between products is not so long (within 1 week).

- As a result, 92,000 product pairs were provided. Then we randomly selected 512 pairs from the individual items, and categorized all the extracted data one-by-one into **"true old and new product pairs"** and **"false (irrelevant) product pairs"** to create *supervised data*.

3. Making Old and New Product Pairs Using Machine Learning Method

3-3. Outline of Characteristics and Classifiers

- When distinguishing true and false of old and new product pairs from the data set, it is necessary to specify labels which serve as indicators (**characteristics**).
- We extracted the following three labels as characteristics:

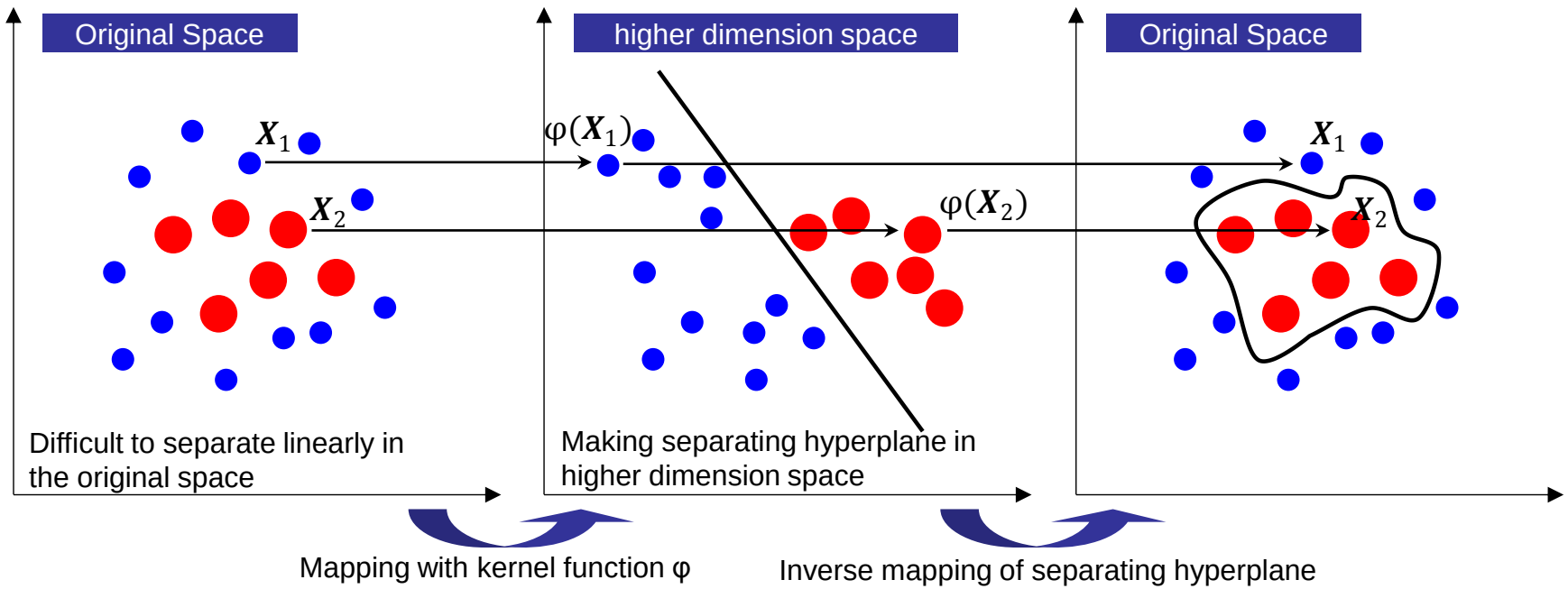
Characteristics used for judgments of old and new product pairs

1. **Jaro-Winkler distance of Product Names**
   Whether the names of paired products are similar.
2. **Zone of product price**
   Whether the paired products belong to the same price zone.
3. **Product launch interval**
   Whether there is a reasonable interval between release dates.

3. Making Old and New Product Pairs Using Machine Learning Method

3-4. Creation of Classifiers

- To implement machine learning, we adopt the **Non-Linear Support Vector Machine** (SVM) as the classifier, coded with Python.



3. Making Old and New Product Pairs Using Machine Learning Method

3-4. Creation of Classifiers (cont'd)

- To improve the classification performance of SVM, we configure (1) the extent to which the complexity of the data boundary surface will be reflected in classifiers (set by **kernel parameters *σ***) and (2) the extent to which faulty identification is allowed (set by **penalty parameters *C***), .
- By using **10-fold cross-validation** and the **grid search method**, we determine hyperparameters (σ, C) which maximize the F-measure to represent performance of classifiers.

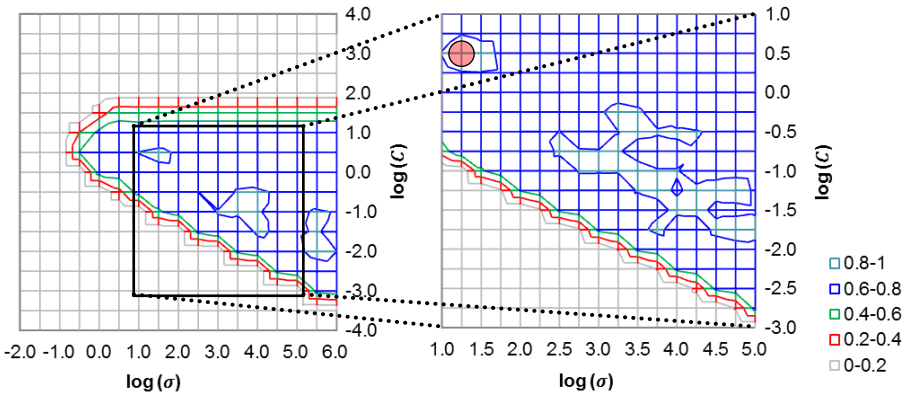
		Actual class	
		P	N
Predicted class	P	True Positive (TP)	False Positive (FP)
	N	False Negative (FN)	True Negative (TN)

$$\text{precision} \equiv \frac{\text{TP}}{\text{TP} + \text{FP}} \qquad \text{recall} \equiv \frac{\text{TP}}{\text{TP} + \text{FN}}$$

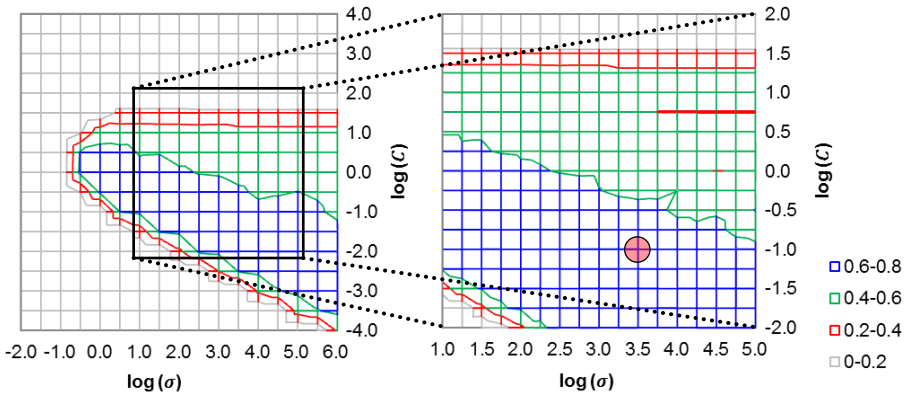
$$\text{F-measure} \equiv \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

Hyperparameters Optimization Using 10-fold Cross-Validation and Grid Search

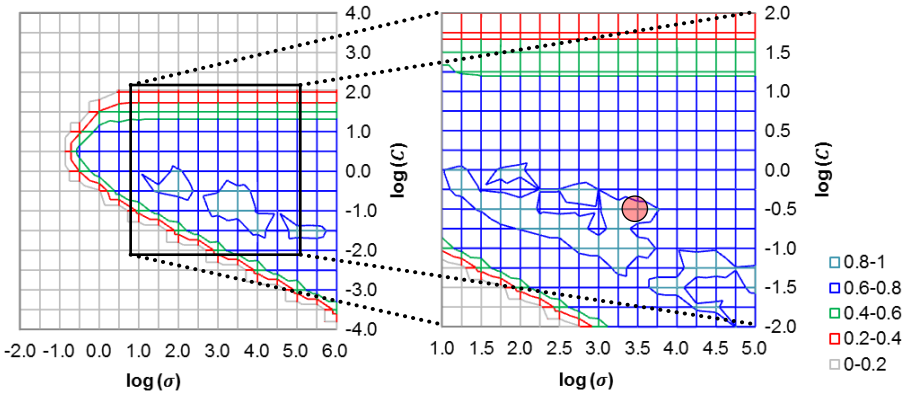
(1) Air conditioners



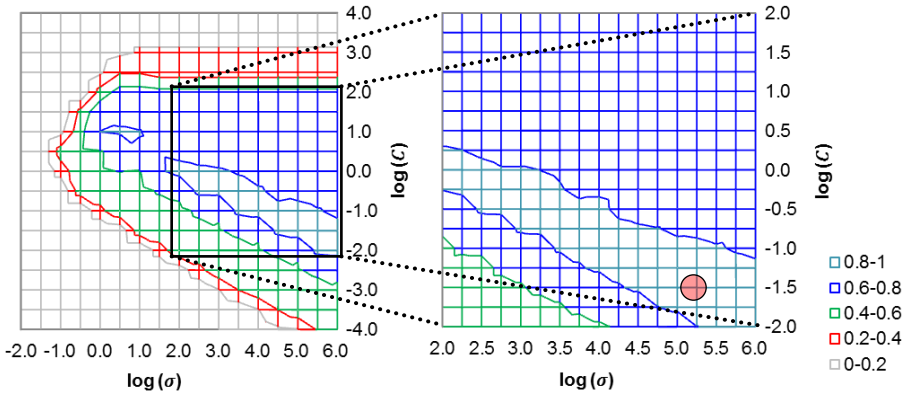
(2) Refrigerators and freezers



(3) Washers and dryers



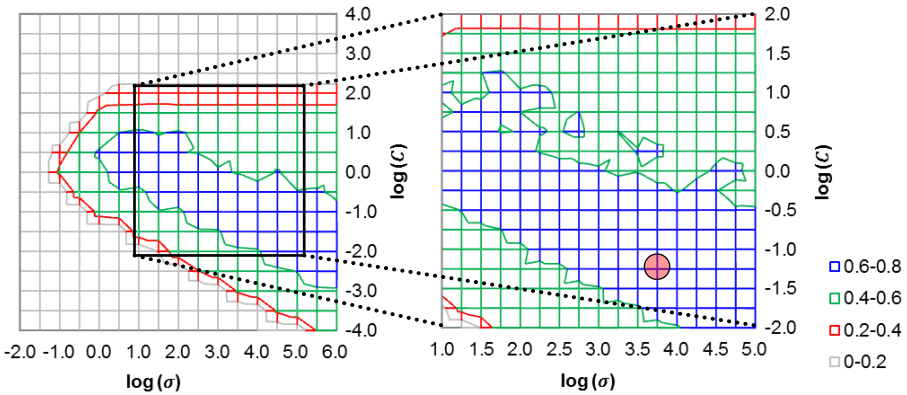
(4) Rice cookers



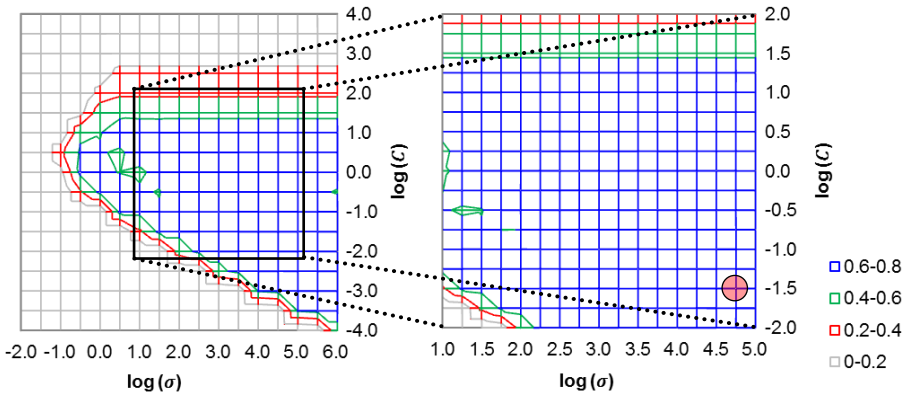
Note: The lattice highlighted in red indicates hyperparameters (σ, C) to maximize F-measure at the time of 10-fold cross-validation obtained by grid search.

Hyperparameters Optimization Using 10-fold Cross-Validation and Grid Search (cont'd)

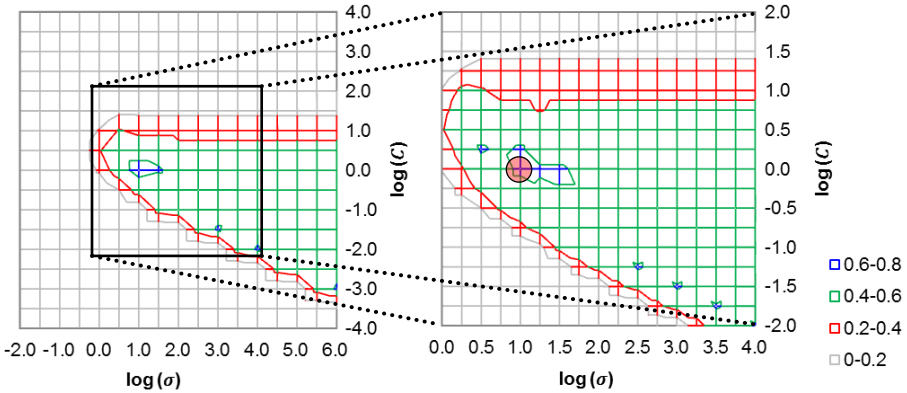
(5) Vacuum cleaners



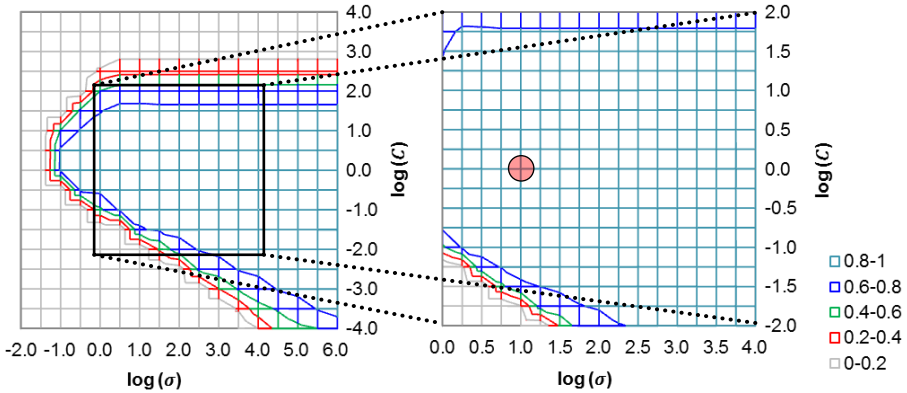
(6) Microwaves



(7) Hair dryers and curling irons



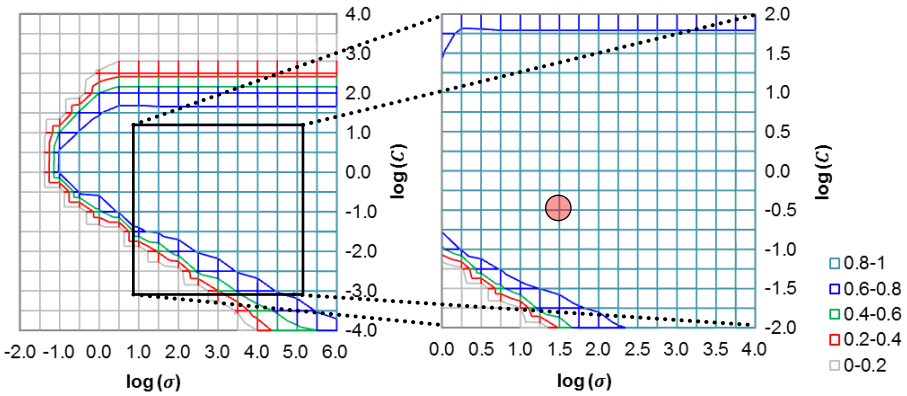
(8) GPS navigations



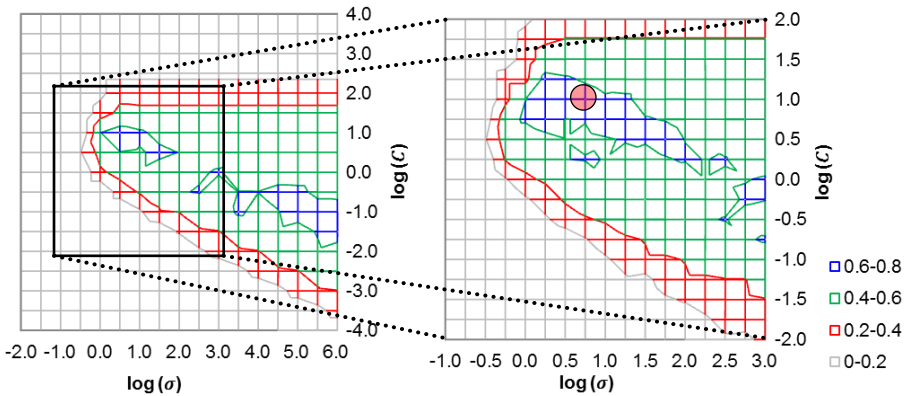
Note: The lattice highlighted in red indicates hyperparameters (σ, C) to maximize F-measure at the time of 10-fold cross-validation obtained by grid search.

Hyperparameters Optimization Using 10-fold Cross-Validation and Grid Search (cont'd)

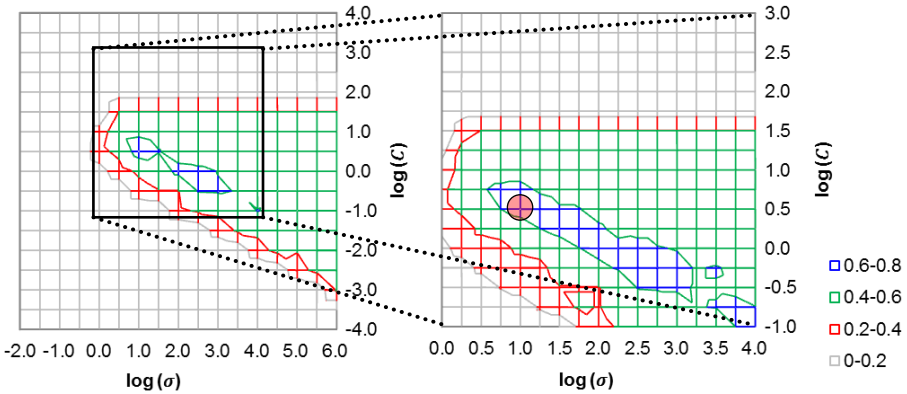
(9) External hard drives



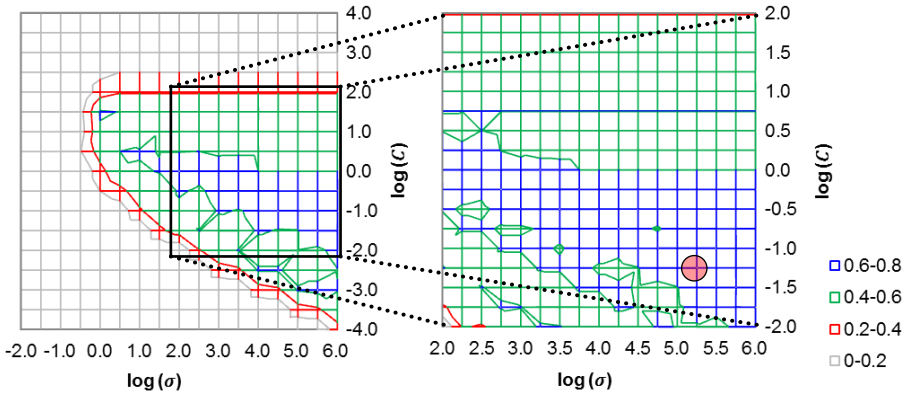
(10) LCD TVs



(11) LCD monitors



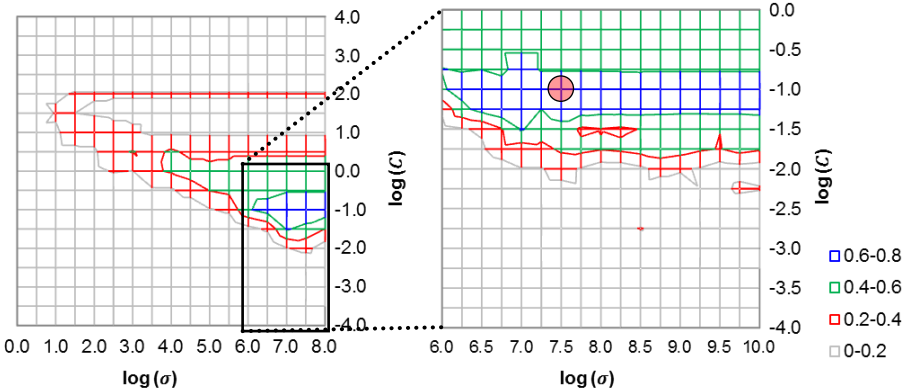
(12) Printers



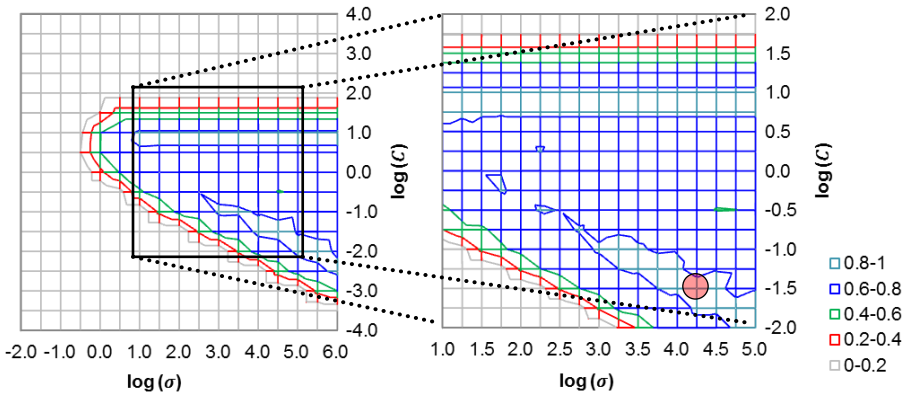
Note: The lattice highlighted in red indicates hyperparameters (σ, C) to maximize F-measure at the time of 10-fold cross-validation obtained by grid search.

Hyperparameters Optimization Using 10-fold Cross-Validation and Grid Search (cont'd)

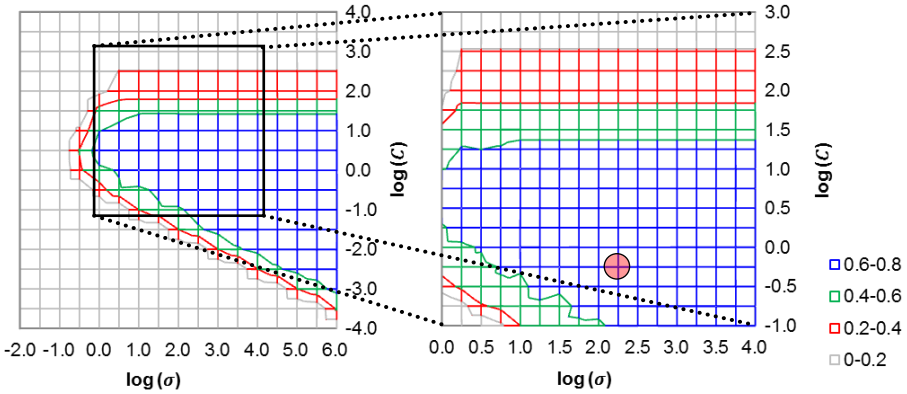
(13) Headphones



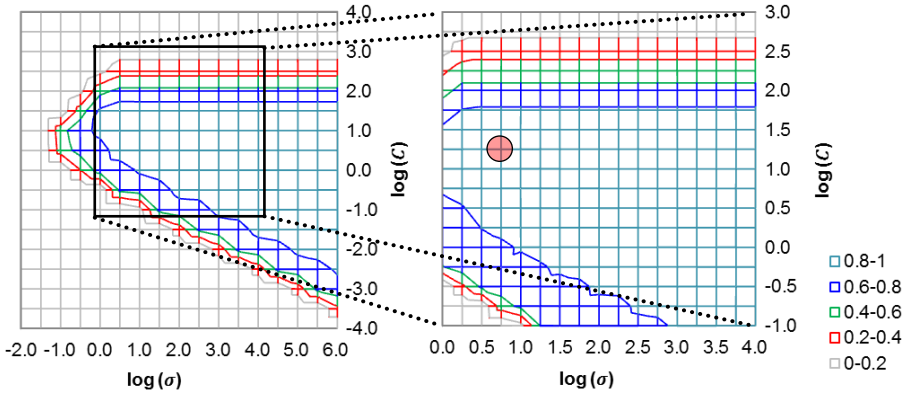
(14) Laptops



(15) Desktops



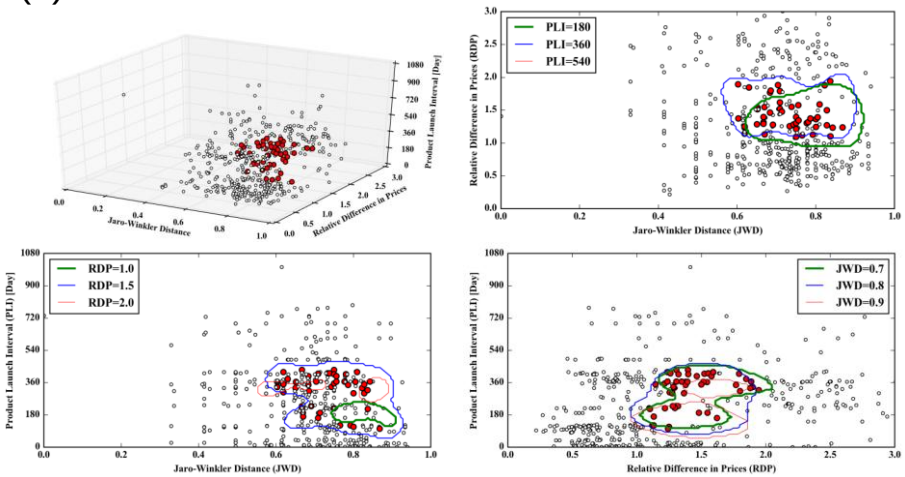
(16) Point-and-shoot cameras



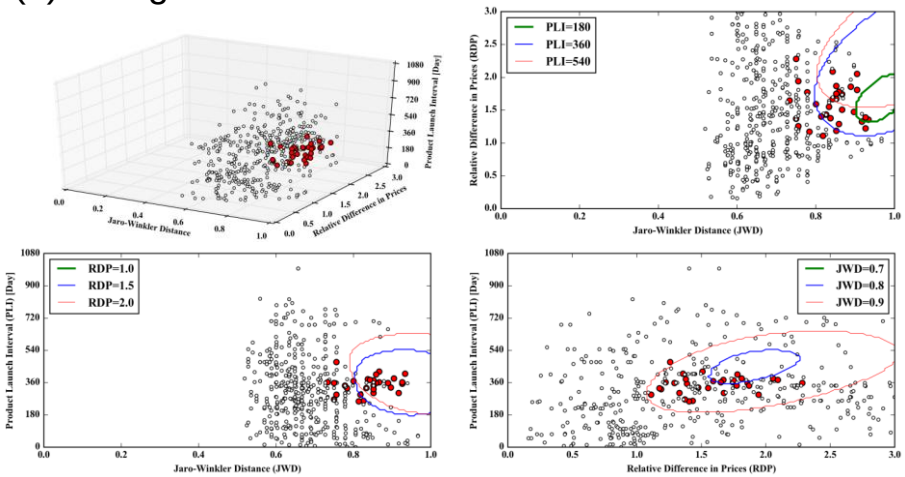
Note: The lattice highlighted in red indicates hyperparameters (σ, C) to maximize F-measure at the time of 10-fold cross-validation obtained by grid search.

Optimal Hyperplanes Using Non-Linear SVM Classifiers

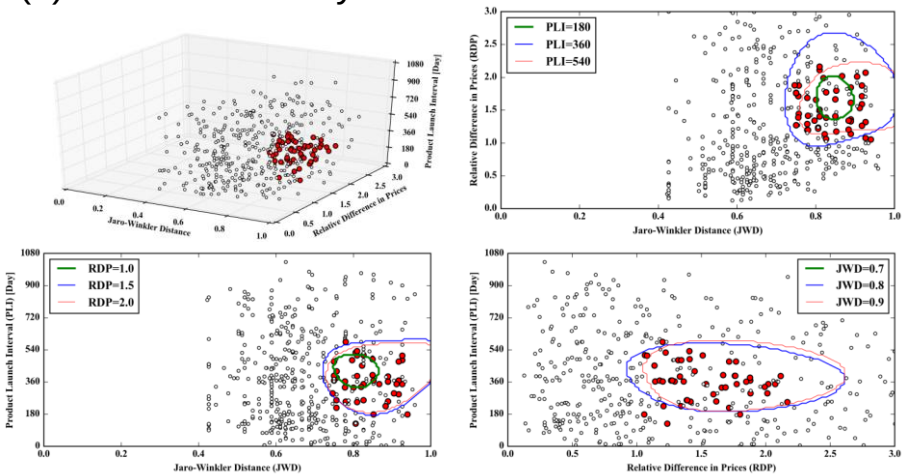
(1) Air conditioners



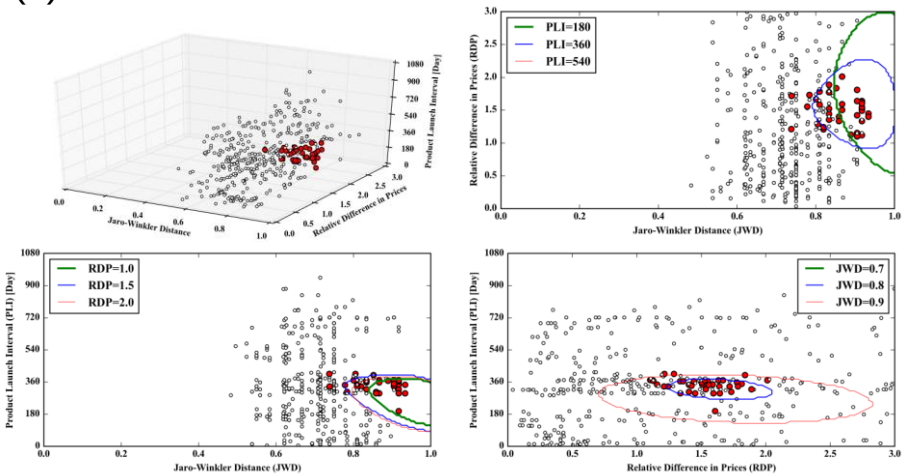
(2) Refrigerators and freezers



(3) Washers and dryers



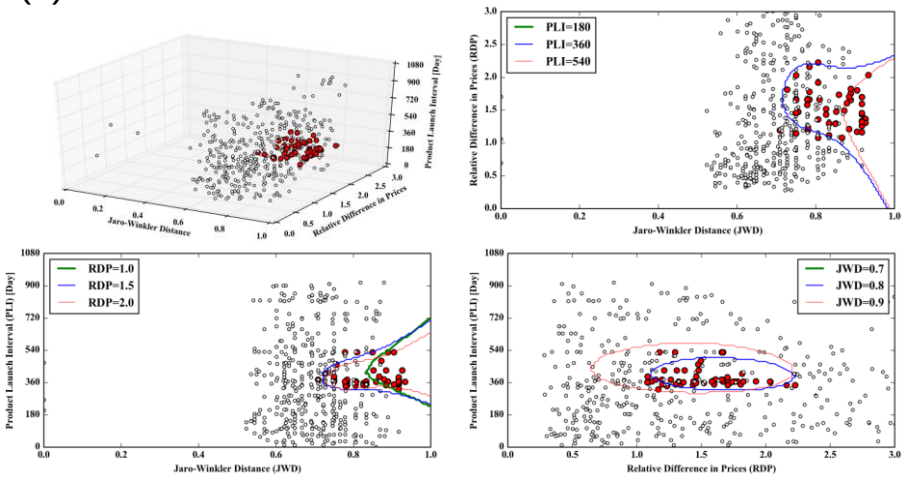
(4) Rice cookers



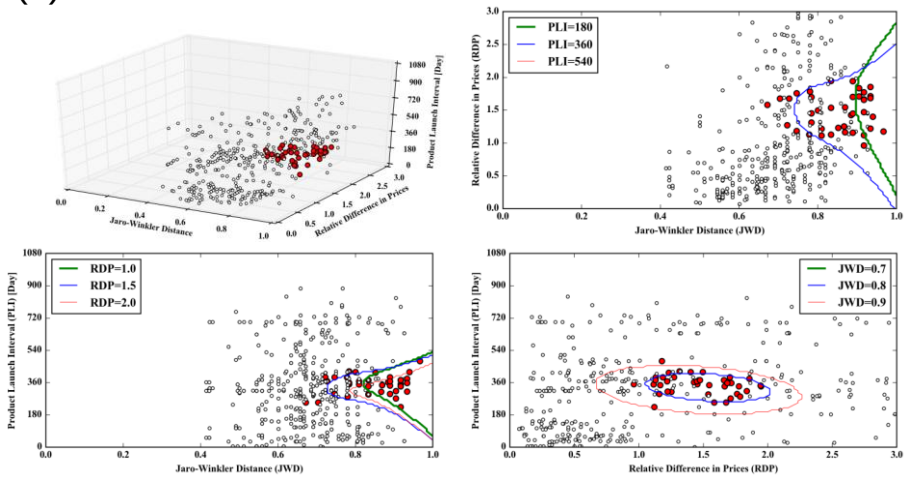
Note: Red dots indicate true old and new product pairs, and white dots represent false product pairs.

Optimal Hyperplanes Using Non-Linear SVM Classifiers (cont'd)

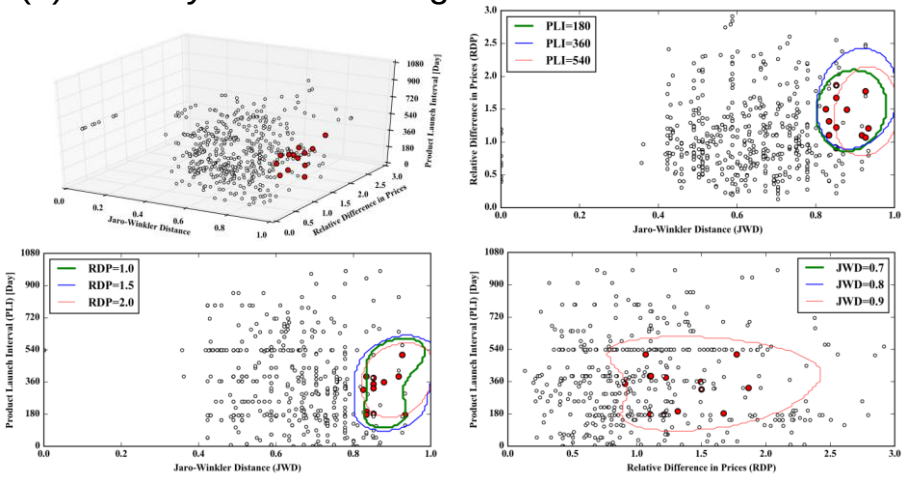
(5) Vacuum cleaners



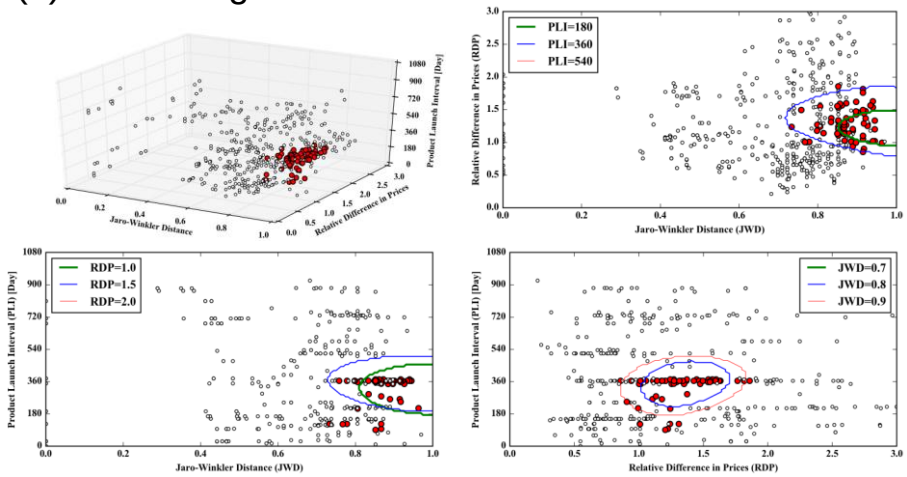
(6) Microwaves



(7) Hair dryers and curling irons



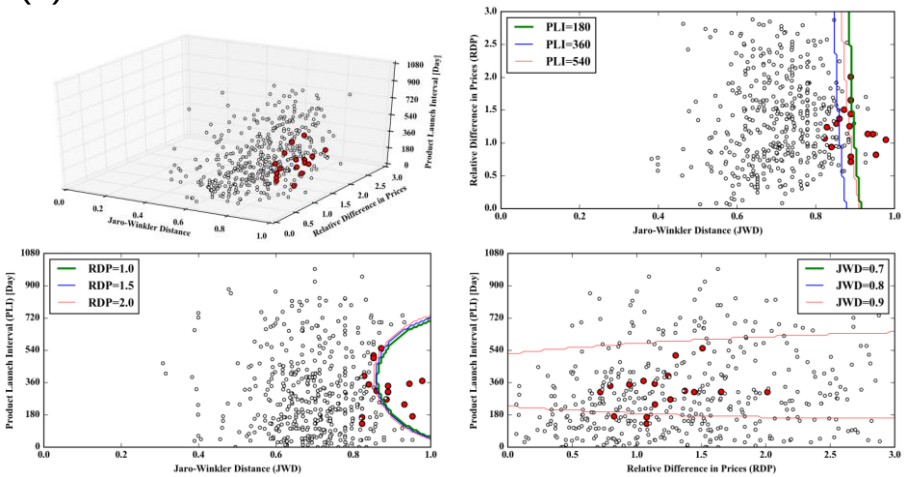
(8) GPS navigations



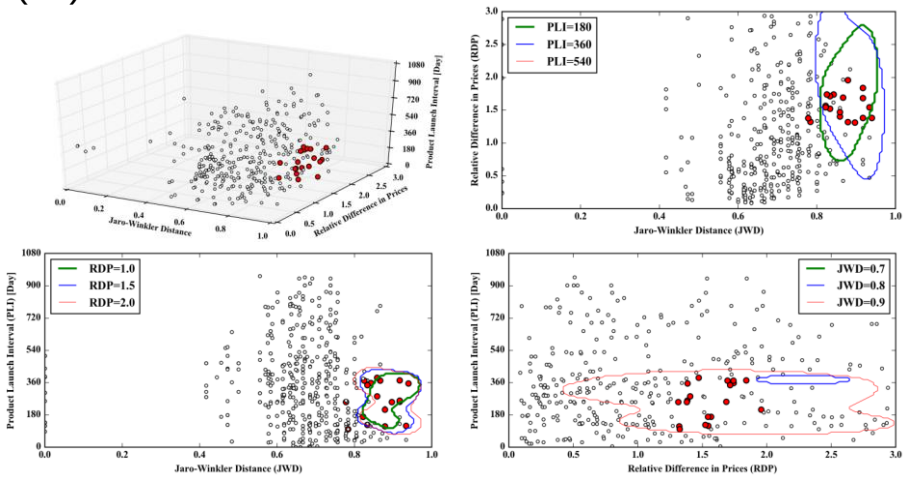
Note: Red dots indicate true old and new product pairs, and white dots represent false product pairs.

Optimal Hyperplanes Using Non-Linear SVM Classifiers (cont'd)

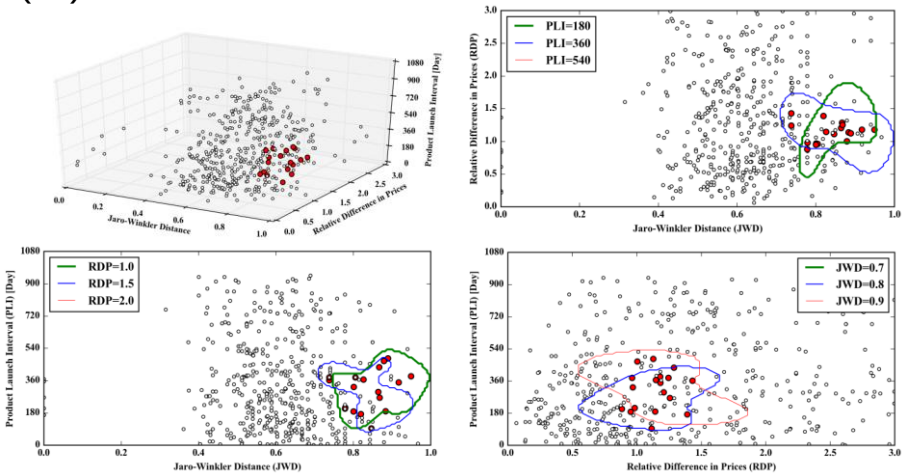
(9) External hard drives



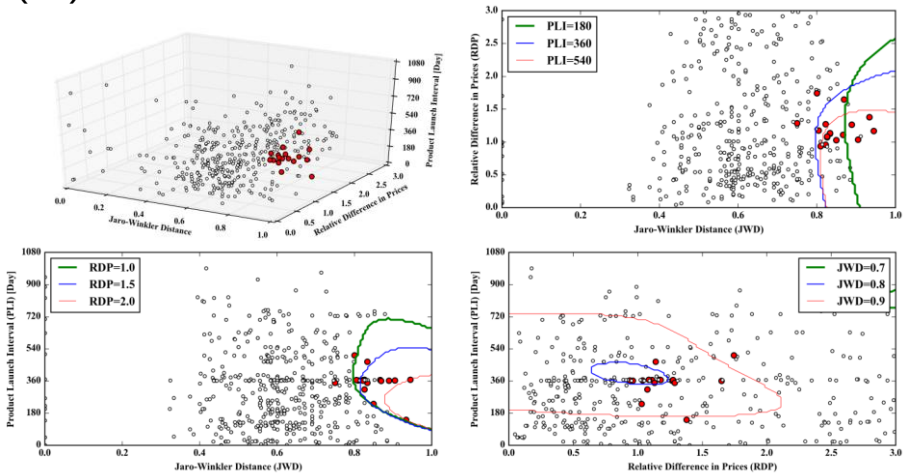
(10) LCD TVs



(11) LCD monitors



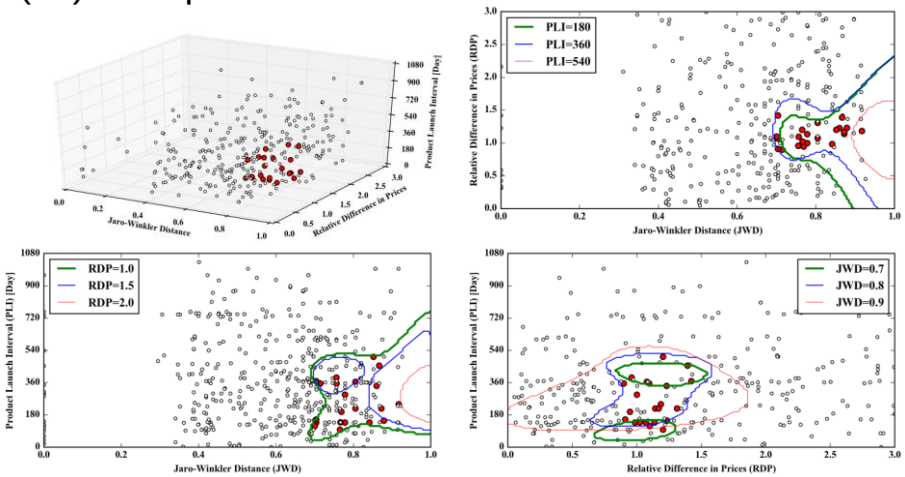
(12) Printers



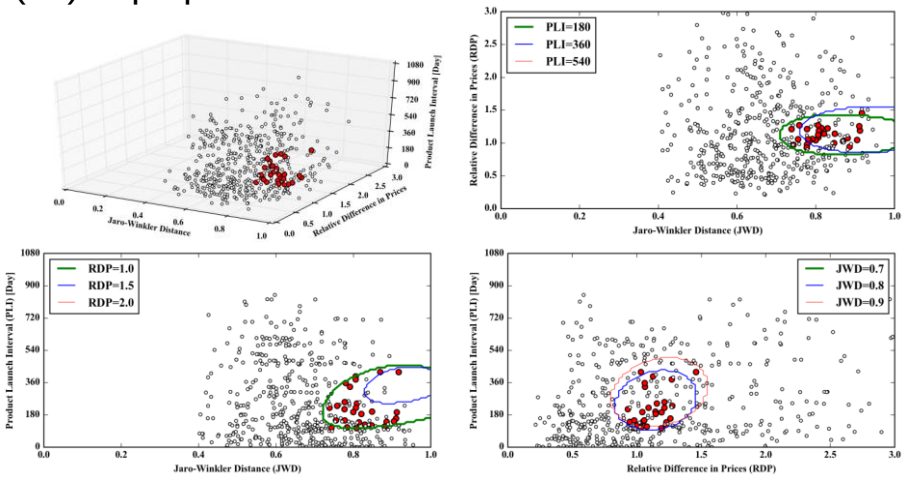
Note: Red dots indicate true old and new product pairs, and white dots represent false product pairs.

Optimal Hyperplanes Using Non-Linear SVM Classifiers (cont'd)

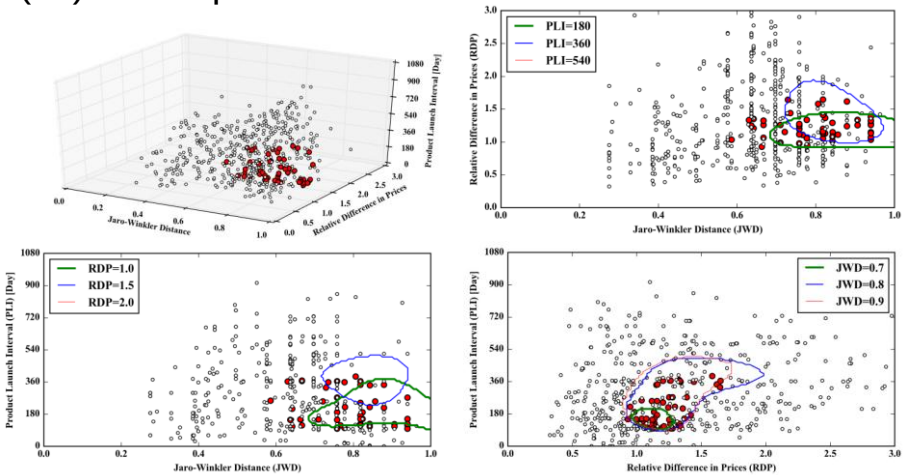
(13) Headphones



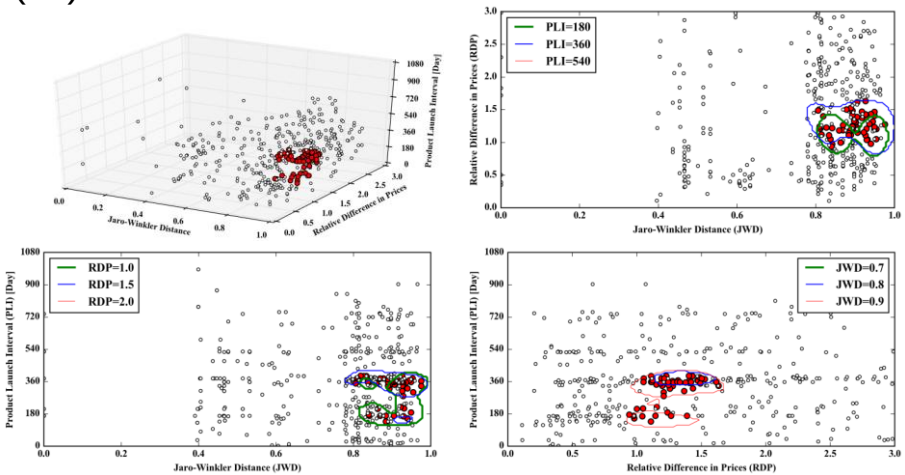
(14) Laptops



(15) Desktops



(16) Point-and-shoot cameras



Note: Red dots indicate true old and new product pairs, and white dots represent false product pairs.

4. Selection of the Quality Adjustment Method and Comparative Analysis of Price Index

4. Selection of the Quality Adjustment Method and Comparative Analysis

4-1. Outline of Major Quality Adjustment Method

- We compile experimental price indexes by applying different quality adjustment methods at the time of product turnover.

1. **Direct Comparison Method** (DCM)
   Method which assumes quality difference between old and new products is ignorable thus processes “price change due to quality changes” as zero.

2. **Overlap Method** (OLM)
   Method which assumes all price difference between old and new products as “price change due to quality changes”.

3. **Matched-Model Method** (MMM)
   Method which calculates the percentage change of price for products which exist in the market in both survey period and following period.

4. Selection of the Quality Adjustment Method and Comparative Analysis

4-1. Outline of Major Quality Adjustment Method (cont'd)

4.

Webscraped prices comparison method (WSM)
Based on results of empirical analysis stating "price change due to quality changes account for approximately 50% of the price differences between old and new products", assumes the portion equivalent to 50% of the webscraped retail price difference as "price change due to quality changes" and the remainder as "pure price change".

- WSM is considered to be less accurate compared to other quality adjustment methods, and at present it is positioned as *second-best method* in the case where other quality adjustment methods are not applicable.

4. Selection of the Quality Adjustment Method and Comparative Analysis

4-2. Hedonic Function Estimation

5.

Hedonic regression method (HRM)
Method which assumes price difference between old and new products are partially due to quality difference arising from product specification. Using large scale data it estimates "price change due to quality changes" by econometric analysis, and processes the remainder as "pure price change."

- Under the same setup as Abe et al. (2016), we estimate the following semi-logarithmic linear hedonic function.

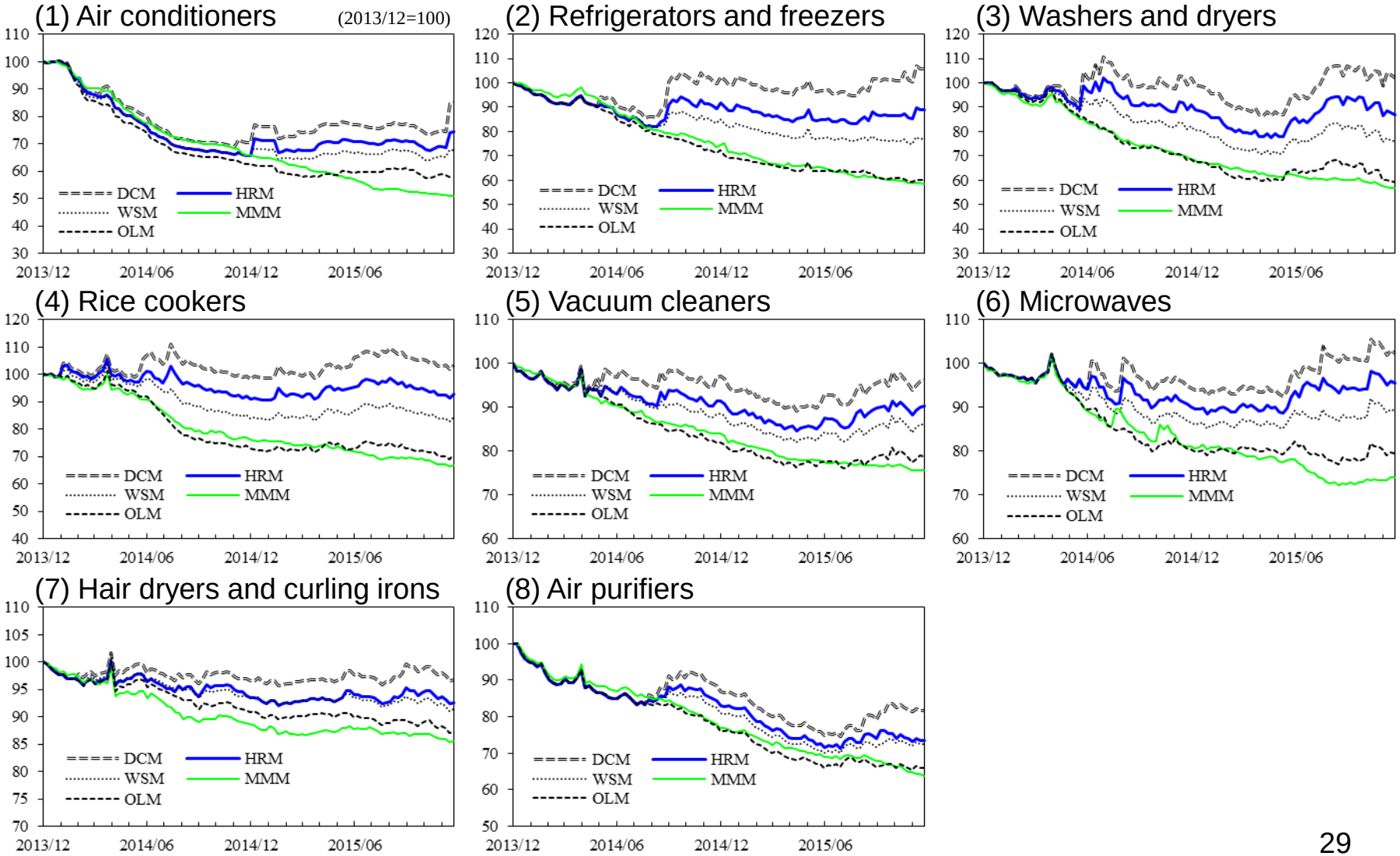
$$\ln(p_{i,t}) = \alpha + \sum_{k} \beta_k X_{i,k} + \sum_{\tau} \gamma_\tau D_t(\tau_i + \tau) + \sum_{\tau} \delta_\tau D_t(\tau) + \varepsilon_{i,t}$$

$$D_t(T) = \begin{cases} 1 & (\text{if} \quad t = T) \\ 0 & (\text{if} \quad t \neq T) \end{cases}$$

$p_{i,t}$: price for product i at time t; $X_{i,k}$: kth specification value of product i
$D_t(\tau_i + \tau)$: dummy variable which controls for the number of weeks elapsed from the release date
$D_t(\tau)$: dummy variable which controls macroeconomic shocks such as price level fluctuations

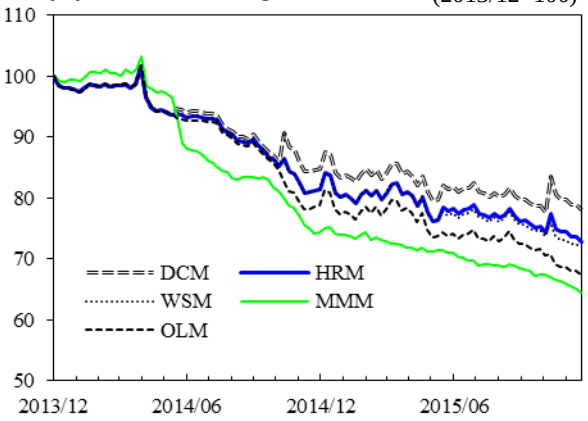
Comparative Analysis of Experimental Price Indexes: Home Electrical Appliances

(2013/12=100)

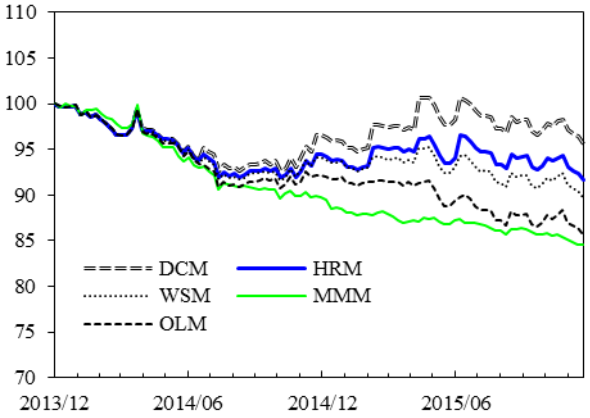


Comparative Analysis of Experimental Price Indexes: Digital Consumer Electronics [1]

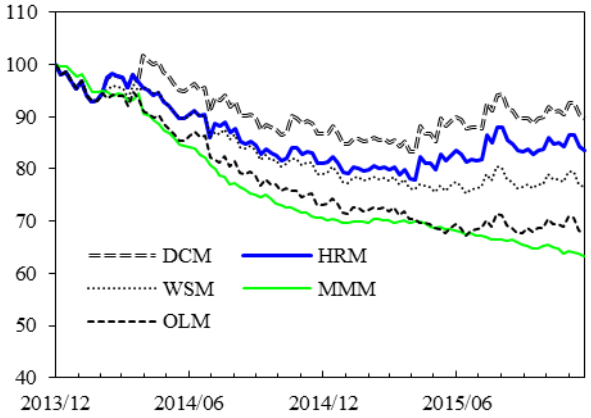
(1) GPS navigations (2013/12=100)



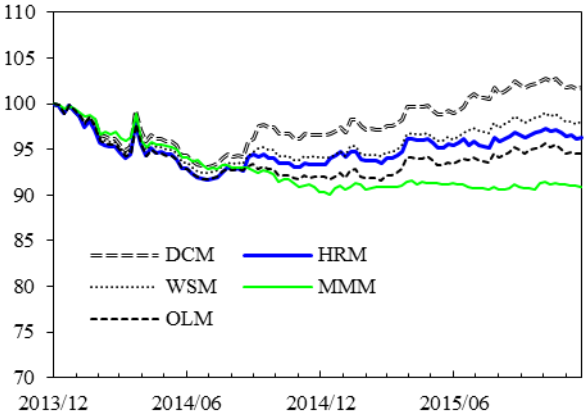
(2) External hard drives



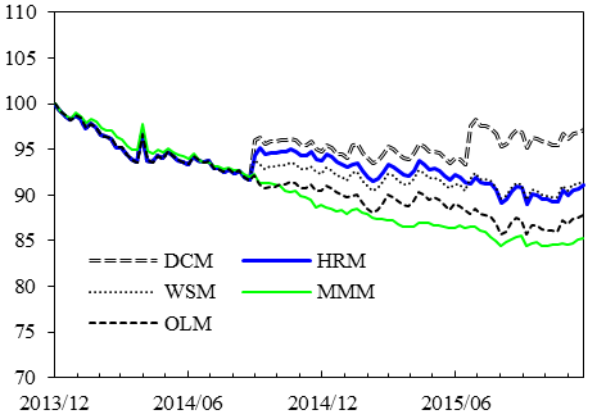
(3) LCD TVs



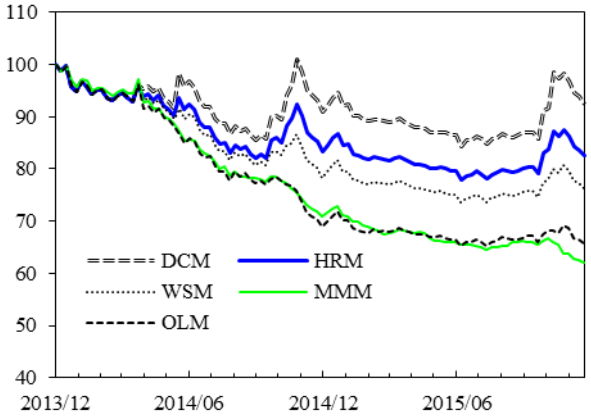
(4) LCD monitors



(5) Printers

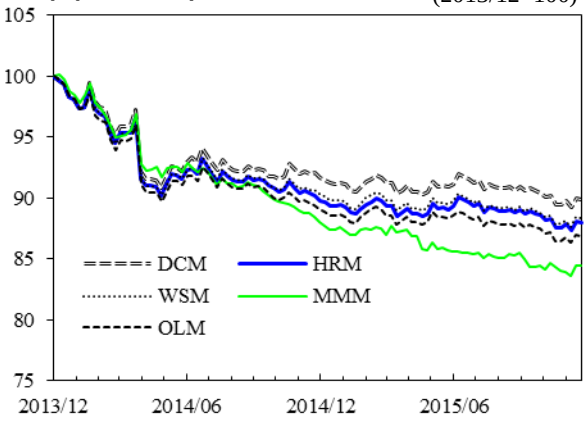


(6) Blu-ray and DVD recorders

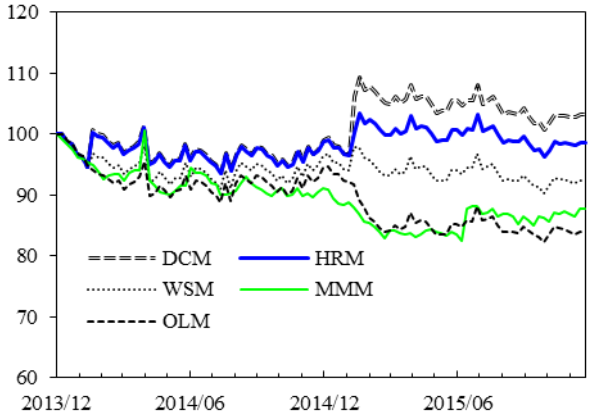


Comparative Analysis of Experimental Price Indexes: Digital Consumer Electronics [2]

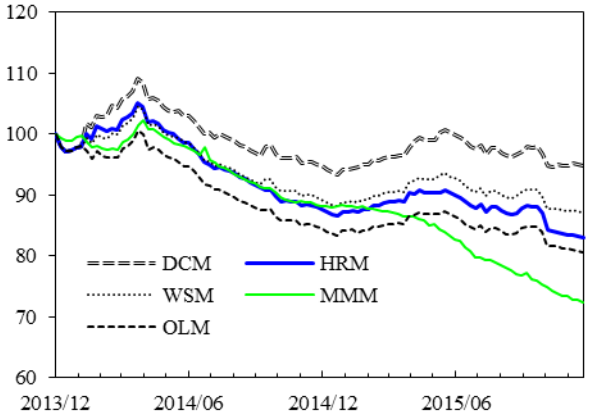
(7) Headphones (2013/12=100)



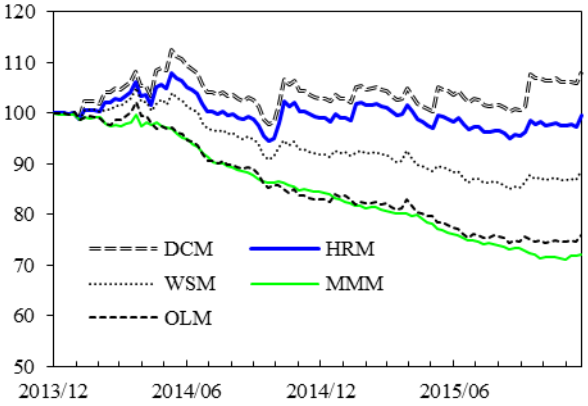
(8) Camcorders



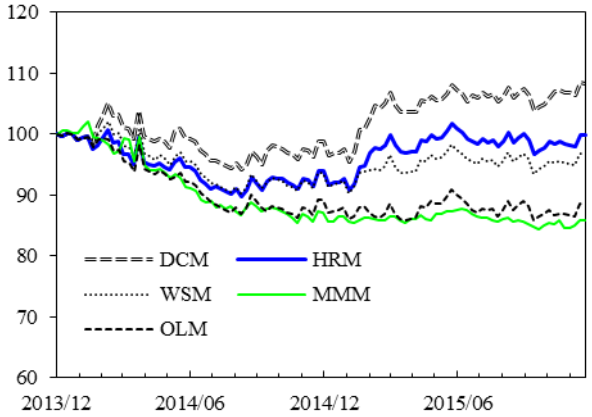
(9) Laptops



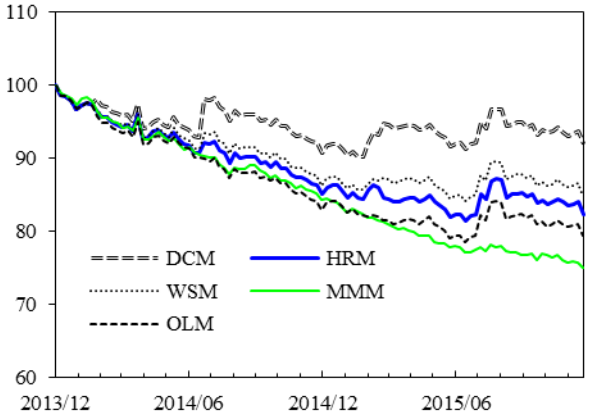
(10) Desktops



(11) Point-and-shoot cameras



(12) DSLR and mirrorless cameras



Comparison of Deviations between Indexes Applied HRM and the Others

	DCM		WSM		MMM		OLM	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Home Electrical Appliances	6.31	5.52	4.02 *	3.29 *	11.96	9.30	11.17	9.52
Air conditioners	4.75	4.21	2.63 *	1.91 *	9.77	7.14	7.58	6.21
Refrigerators and freezers	9.70	8.08	6.28 *	4.90 *	16.56	13.46	17.02	13.72
Washers and dryers	8.85	7.72	7.44 *	6.46 *	19.80	17.12	18.54	16.23
Rice cookers	7.65	6.84	6.84 *	6.08 *	17.88	15.87	17.54	15.73
Vacuum cleaners	4.42	3.97	2.54 *	2.11 *	7.64	6.43	7.61	6.48
Microwaves	4.25	3.53	4.09 *	3.41 *	12.47	9.96	10.35	8.73
Hair dryers and curling irons	2.97	2.70	0.63 *	0.46 *	5.13	4.57	3.07	2.55
Air purifiers	3.81	2.89	1.45 *	1.12 *	4.69	3.90	5.24	4.16
Digital Consumer Electronics	4.88	4.37	0.96 *	0.85 *	7.15	5.93	5.48	4.97
GPS navigations	2.83	2.23	0.44 *	0.23 *	6.39	5.86	2.79	2.12
External hard drives	2.52	1.92	1.11 *	0.76 *	5.36	4.27	3.60	2.66
LCD TVs	5.21	4.76	3.83 *	2.82 *	11.94	10.45	9.65	8.26
LCD monitors	3.38	2.93	0.94 *	0.81 *	3.40	2.84	1.40	1.13
Printers	3.08	2.00	0.80 *	0.58 *	4.09	3.36	2.67	2.11
Blu-ray and DVD recorders	6.07	5.20	3.87 *	3.20 *	11.56	9.68	11.14	9.42
Headphones	1.45	1.30	0.29 *	0.22 *	2.33	1.87	0.76	0.71
Camcorders	3.25 *	2.29 *	4.62	4.11	10.39	8.91	10.49	8.96
Laptops	7.27	6.62	1.93 *	1.59 *	5.25	3.65	3.46	3.32
Desktops	4.51 *	4.02 *	7.32	6.30	16.73	14.75	15.73	13.94
Point-and-shoot cameras	5.67	5.29	2.17 *	1.73 *	8.74	7.20	7.58	6.20
DSLR and mirrorless cameras	6.99	6.06	1.75 *	1.45 *	4.22	3.12	2.29	2.04

We conduct periodic averaging using **RMSE** (Root Mean Squared Error) and **MAE** (Mean Absolute Error) on the deviation between indexes applied HRM and the other quality adjustment methods.

The numbers with asterisk imply the smallest deviation from the results of HRM.

5. Final Remarks

5. Final Remarks

- In this paper, we combine features of both ***traditional* approach** and ***non-traditional* approach** by applying machine learning methods in order to pair old and new products accurately. Based on the results, we show the appropriateness of **Webscraped Prices Comparison Method** (WSM).
- Also, we verify that applying **Matched-Model Method** (MMM) to products pushing back price frequently at product turnover may bring about a downward bias.
- The contribution of this paper is to present a new price index compilation method by using big data, and once again indicating the possibility of potential bias.