Discussion Paper Series A No.722

Lower-Level Substitution Bias in the Japanese Consumer Price Index: Evidence from Government Micro Data

Shigenori Shiratsuka
(Keio University)

June 2021

Institute of Economic Research
Hitotsubashi University
Kunitachi, Tokyo, 186-8603 Japan

Lower-Level Substitution Bias in the Japanese Consumer Price Index: Evidence from Government Micro Data *

Shigenori Shiratsuka†

June 16, 2021

Abstract

This paper explores measurement errors in the Japanese Consumer Price Index (CPI) stemming from lower-level substitution within items. The CPI is widely used as a measure for inflation or the cost of living. The Japanese CPI employs the one-specification for one-item policy in surveying individual prices. The policy specifies a few most popular specifications for each item and continuously surveys their prices at specific outlets. As a result, the price homogeneity is generally maintained, limiting the impact of the differences in the elementary aggregation formulas, which corresponds to the narrow definition of the lower-level substitution bias. In contrast, the price representativeness becomes difficult to be maintained for highly heterogeneous and differentiated products. That is another aspect of the lower-level substitution bias particular to the Japanese CPI, encompassed by the broad definition of the lower-level substitution bias. However, quantitative assessments on the lower-level substitution bias in the Japanese CPI are very limited since the detailed CPI data at individual price observations was not readily available for a long time. This paper is the first trial on a quantitative assessment of the lower-level substitution bias using the micro data for the Retail Price Survey (RPS), which is the primary source data for the Japanese CPI. Empirical evidence confirms that the lower-level substitution bias in the Japanese CPI differs from that for the U.S. CPI. On the one hand, the one-specification for one-item policy in the price survey succeeds in keeping price observations homogeneous, limiting the elementary aggregation bias. On the other hand, the policy also weakens price representativeness, which requires additional quantitative assessments using alternative data sources, such as scanner data and web-scraping data.

Keywords: Consumer Price Index; Measurement Errors; Substitution Effects; Elementary Aggregation Formula; Price Representativeness; Price Survey Method

JEL codes: C43; E31

*This paper was prepared as a part of the FY2020 research project entitled "An Economic Analysis for Consumer Prices using Government Micro Data" under the Joint Usage and Research Center, Institute of Economic Research, Hitotsubashi University. Micro data for the Retail Price Survey used in this paper was provided by the Statistic Bureau of Japan, based on Article 33 of the Statistics Act. The author thanks Erwin Diewert, Shin'ichi Fukuda, Masahiro Higo, Yukinobu Kitamura, Nao Sudo, Tsutomu Watanabe, and seminar participants at Hitotsubashi University and TCER Financial Section for their constructive comments and discussions. The author also thanks Masayuki Tanaka for his assistance. The author acknowledges the financial support from Joint Usage and Research Center, Institute of Economic Research, Hitotsubashi University, and Zengin Foundation for Studies on Economics and Finance.

†Keio University. E-mail: shigenori.shiratsuka@keio.jp

1 Introduction

In this paper, I explore measurement errors in the Japanese Consumer Price Index (CPI) stemming from lower-level substitution within items.

The CPI is widely used as a measure of inflation or the cost of living by generally using a fixed-weight Laspeyres formula. This formula is used mainly because of its ease of calculation and comprehension, limiting the total cost of constructing the statistics. However, such simplicity makes it difficult for the CPI to account for the dynamic nature of the economic activity, such as changes in consumers' behavior in response to relative price fluctuations between goods and services, the introduction of new goods, and the disappearance of old goods. As a result, measurement errors are introduced into the CPI.

Although the measurement problems for the CPI attract considerable attention from practitioners and academics, Japan has very limited research on assessing the CPI measurement errors compared with the United States, as summarized in Table 1. My previous study, Shiratsuka (1999), is the first trial to make an overall assessment of measurement errors in the Japanese CPI, based on the best available information. However, given such limited previous studies, the estimate needs various rather bold assumptions. Since then, the situation has almost unchanged over 20 years, as seen in the very limited follow-up studies to assess the overall measurement errors in the Japanese CPI. That is a very contrasting situation in the United States as we observe many following-up studies on the so-called *Boskin Report*, Boskin et al. (1996).[1]

Among various factors for the measurement bias of the CPI, the lower-level substitution bias has the least previous studies with rigorous quantitative assessment in Japan, since the data on the CPI lower than the item level has not been readily available for a long time. Without micro data for the CPI at an individual price level, it is impossible to examine the elementary aggregation problem from individual price observations to item index. This paper is the first trial on a quantitative assessment of the lower-level substitution bias using the micro data for the Retail Price Survey (RPS), which is the source data for the Japanese CPI.

The substitution bias arises in the CPI because consumers change their purchasing behavior in response to relative price changes. This substitution bias can be divided into the low-level and upper-level substitution biases, corresponding to the two-step procedure in the CPI compilation. In the first step, the elementary aggregation compiles the price index for the item level at each region from individual price observations at outlets.[2] That is the lowest level of groups of goods and services surveyed in the CPI, which is supposed to include relatively homogeneous

[1] The winter issue in 1998 of The Journal of Economic Perspectives features the symposium entitled "Measuring the CPI," which include the summary of the so-called Boskin Report, Boskin et al. (1998), the responses from staff at the Bureau of Labor Statistics, Abraham et al. (1998), and other related issues by major economists in this field, such as Deaton (1998), Diewert (1998), Nordhaus (1998), and Pollak (1998). In addition, there exists a huge list of following-up studies on Boskin et al. (1996), such as General Accounting Office (2000) and Lebow and Rudd (2003). The Bureau of Labor Statics staff makes additional public responses to such studies in Greenlees (2006). In contrast, there are very few studies on the CPI measurement issues, including Ariga and Matsui (2003).

[2] For the details on the elementary aggregation index formulas, see Diewert (2021) and International Monetary Fund (2020).

Table 1: Measurement Errors in CPI

	Japan			United States		
	Shiratsuka (1999)	Shiratsuka (2006)	Broda & Weinstein (2007)	Boskin Report (1996)	Lebow & Rudd (2 003)	Broda & Weinstein (2007)
Upper-level substitution	0.00	→	0.2	0.15	0.30	0.0
Lower-level substitution	0.10	→	0.4	0.25	0.05	0.0
Quality changes / new products	0.70	↓	0.8 + 0.2	0.60	0.37	0.8
Outlet substitution	0.10	↘	0.2	0.10	0.05	0.2
Weighting	—	—	—	—	0.10	—
Total	0.90	↓	1.8	1.10	0.87	1.0

goods and services with similar characteristics, content, price, and price fluctuation. Second, those elementary price indexes are aggregated into higher-level price indexes by employing respecified index formula, basically Laspeyres price index.

The Japanese CPI was thought to be less influenced by the lower-level substitution bias than the U.S. CPI due to their differences in price survey methods. The Japanese CPI employs the one-specification for one-item policy in surveying individual prices. The policy specifies a few most popular specifications for each item and continuously surveys their prices at specific outlets. As a result, the dispersion of price observations is generally restricted at a low level, limiting the impact of the differences in elementary aggregation formulas for aggregating individual price observations into item indexes. However, the price representativeness becomes difficult to maintain for highly differentiated products and services.

In contrast, the U.S. CPI employs random sampling of many price observations within broadly defined items. This sampling method becomes a source of the lower-level substitution bias when using an arithmetic mean of price observations as an elementary aggregation formula at the lower level. The random sampling method effectively maintains the price representativeness while ineffectively controls the price homogeneity.

The U.S. CPI changed its elementary aggregation formula from the arithmetic mean to the geometric mean in about two-third of the item strata in 1999. In that sense, the lower-level substitution bias stemming from the elementary aggregation formula is unique to the United States because of the random sampling of a large number of price observations. At least, this problem is less relevant to Japan. Instead, Japan's problem arises mainly from the one-specification for one-item policy by weakening the price representativeness. I call this issue the broadly-defined lower-level substitution bias, encompassing the elementary aggregation bias as the narrowly-defined lower-level substitution bias.

With the first trial to employ the micro data for the RPS, I confirm that the lower-level substitution bias in the Japanese CPI differs from that examined by Boskin et al. (1996) for the U.S. CPI. On the one hand, the Japanese CPI successfully constrains the elementary aggregation bias, or the narrowly-defined lower-level substitution bias, at a smaller margin than the U.S. CPI, reflecting the more homogeneous price observations within the items with the one-specification for one-item policy. On the other hand, the Japanese CPI has problems of the broadly-defined lower-level substitution bias, which corresponds to the weakened price representativeness due to the one-specification for one-item policy in the price survey. Some items with the loosely defined specification have features of highly heterogeneous and differentiated products within the item, resulting in the intensified price representative problems, in addition to the elementary aggregation problem.

I should emphasize two things in making a quantitative assessment on the CPI measurement errors. First, the measurement of individual prices raises more difficult and fundamental problems in examining price measurement than the aggregation of those prices into the overall price index. The problems of aggregation are well understood by economists, and workable solutions are now readily available to some extent, including the introduction of the chain-weight index formula. However, collecting individual prices with high accuracy is fairly difficult than generally thought since it is impossible to survey prices for all goods and services traded in the economy. In addition, the nominal values of goods and services are difficult to be decomposed into quantity and price on a quality-adjusted basis. Second, the assessment of the measurement errors in the CPI needs to consider practical differences in CPI compilation methods from country to country, even though the measurement biases are examined with a common and general framework.[3]In that respect, the most significant differences across countries are observed in collecting individual prices for goods and services.

This paper is structured as follows. Section 2 explains RPS as the primary source data for the CPI in Japan with focusing on its particular feature of the one-specification for one-item policy in the price survey. Section 3 examines the practice of price substitutions in the Japanese CPI by estimating the implied quality index from the CPI and RPS. Section 4 summarizes the major elementary aggregation formula and estimates the elementary aggregation bias, which corresponds to the narrowly-defined lower-level substitution bias. Section 5 carries out a case study analysis for selected items to examine the price representativeness issues. Finally, Section 6 concludes the paper by assessing the lower-level substitution bias in the Japanese CPI and providing some proposals to improve its precision.

[3] See Shiratsuka (2007) for the counterargument against Broda and Weinstein (2007), which naively apply U.S. empirical results to Japan, without considering the differences between the two countries, such as economic and price developments and CPI compilation methods.

2 Retail Price Survey and Data Employed

2.1 Overview of Retail Price Survey

The RPS surveys retail prices of goods and services selected based on the importance of household consumption expenditure.[4] The RPS is a primary data source for the CPI in Japan.

The price survey is structured by three types, as summarized in Table 2. First, enumerators survey relatively standardized products and services purchased at outlets nearby, categorized into four groups: Item group A covers goods and services consumers usually purchase in their residential districts, and the prices vary by district, such as food and domestic non-durables; Item group B covers goods and services consumers usually purchase in their representative commercial areas or large retail stores with price differentials across stores, such as clothing and electric appliances; Item group C covers goods and services with small price differentials across stores, including recreational goods; Item group S includes other goods and services that are surveyed without specifying survey district, such as contact lenses, and gasoline.

Second, the prefectural governments survey goods and services with almost uniform prices in each prefecture or city, town, and village, such as water charges, hospital charges for delivery (item group D). Third, the Statistics Bureau surveys goods and services with uniform prices throughout the country or within a region, such as electricity and telephone charges (item group E).

Table 2: Price Survey Framework in the Retail Price Survey

Survey Group	Surveyor	Types of Products	Items	Number of Prices Surveyed
A	enumerator	Shopped nearby	Food, Daily necessities, ...	42
B	enumerator	Shopped at large retailers	Clothing, Electric appliances, ...	21
C	enumerator	Small price differences at outlets	Recreational goods, ...	12
D	Prefectural Government	Utility charges at local level	Water charges, ...	—
E	Statistics Bureau	Utility charges at nation-wide level	Electricity, Telephone charges, ...	—
S	enumerator	No specific survey districts	Gasoline, ...	

Source: Statitics Bureau of Japan, "Outline of the Retail Price Survey" (https://www.stat.go.jp/english/data/kouri/doukou/pdf/outline.pdf)

[4] In January 2013, "Trend Survey" part was added to the RPS to explore the price structure by region and store type.

2.2 Data employed

To examine the lower-level substitution bias in the Japanese CPI, I narrow the focus in two respects among the all available micro data provided by the Statistic Bureau of Japan (see Table 3): First, only data for Tokyo Metropolitan Area, which is commonly available both the CPI and the RPS at an item level; and data for the period from January 2015 to December 2019, corresponding to the 2015-base CPI.

Table 3: Retail Price Survey as a Source Data for the Consumer Pice Index

	Number of items			Weight		
	Number of items	Share		weight	Share	
		in total	in category		in total	in category
CPI (Tokyo Metropolitan)	582			10,000		
Total	462	79.4%	79.4%	4,392	43.9%	43.9%
Goods	402	69.1%	91.2%	3,287	32.9%	76.1%
Every 10-day survey	44	7.6%	97.8%	332	3.3%	99.6%
Services	60	10.3%	42.6%	1,105	11.0%	19.4%

The table shows that the data used in this paper covers 79.4 percent of the CPI in term of the number of items, and 43.9 percent in terms of the weight, since the RPS data does not include the CPI items with large weight, such as house rent, public transportation fares, school fees, and medical services charges. The data also covers 91.2 percent of goods surveyed, while just 42.6 percent of services surveyed in terms of the number of items. The coverage declines to 76.1 percent for goods and 19.4 percent for services in terms of the weight. Among the goods surveyed, some 44 items are surveyed every ten days since some goods with high price volatility require a more frequent price survey than other standard goods, to trace price movements.

2.3 One-specification for one-item policy

In exploring the lower-level substitution bias in the Japanese CPI, it should be noted that sources of the bias are not restricted for the elementary aggregation formula.

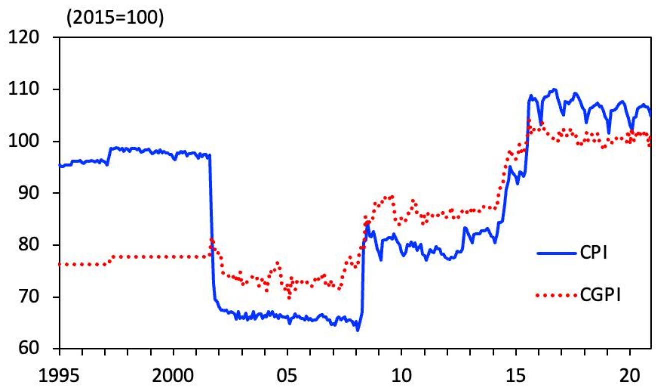
The Japanese CPI employs the one-specification for one-item policy in surveying individual prices. This method specifies a few most popular specifications for each item and continuously surveys its prices at some specific outlets.

The policy has an advantage in that monitoring the specification changes for the items surveyed, including product downsizing/upsizing, becomes easier. The policy, however, has a disadvantage in that the specified item is not necessarily representative of all items in an expenditure category. The tightly defined specification can sometimes introduce a significant noise when there are major changes in the pricing strategy of the products.

Figure 1 plots the price indexes for chocolate in the CPI and the CGPI (Corporate Goods

Price Index).[5]The CGPI surveys average prices for readily available some specific products for chocolates, while the CPI makes a very detailed specification of products by listing three product names: "Meiji milk chocolate," "Lotte Ghana milk chocolate," or "Morinaga milk chocolate." The figure shows a significant downward shift in 2001 when major specification changes occurred for these specifications. Just ignoring the sharp downward shift in the CPI caused by these changes and linking the trend before and after the specification changes, the CPI movements are very close to those in the CGPI. The figure also shows two sharp upward shifts in the CPI in 2008 and 2015, while the upward shifts are much mild and gradual in the CGPI.

Figure 1: CPI and CGPI for Chocolate



Sources: Statistics Bureau of Japan, "Consumer Price Index"; Bank of Japan, "Corporate Goods Price Index."

The price survey practice in the Japanese CPI seems to constrain the lower-level substitution bias stemming from the elementary aggregation formula by keeping price observations homogeneous. It likely induces another lower-level substitution bias due to the failure in representing the overall price movements for the expenditure categories.

3 Price Substitution Practice

As a preparation for the quantitative analysis on the lower-level substitution bias, I examine price substitution practice in the CPI compilation procedure through simulation analysis. The simulation clarifies the price linkage between the RPS and the CPI in a practical manner.

3.1 Price Substitution Simulations

The CPI is computed from the RPS (or unit prices in general) by making a quality adjustment when changing the item specifications. Quality adjustment covers various aspects of product characteristics, such as product performances and unit size. The CPI is computed from the RPS with taking account of quality changes at price substitution by applying the equation below.

[5]The CGPI covers goods traded in the corporate sector, compiled by the Bank of Japan. The prices of products are continuously surveyed by keeping the quality and contract terms fixed.

$$CPI_{i,t} = \frac{RPS_{i,t}}{QAL_{i,t}}, \tag{1}$$

where $CPI_{i,t}$, $RPS_{i,t}$, and $QAL_{i,t}$ denote the CPI, the RPS, and implied quality index of item i at period t with normalized at base period 0. Thus implied quality index is computed from the CPI and the RPS, although details of the quality adjustment at an individual price level are not readily available.

$$QAL_{i,t} = \frac{\frac{RPS_{i,t}}{RPS_{i,0}}}{CPI_{i,t}} \tag{2}$$

The Statistics Bureau of Japan makes an announcement on price substitutions when the item specification changes. The announced price substitutions are common to all price observations in the item. In practice, however, the Statistics Bureau of Japan continues to make various adjustments to link discontinued price observations to compile the CPI.

Using equation (2), I compute the monthly quality adjustment implied by the CPI and the RPS as follows.

$$\frac{QAL_{i,t}}{QAL_{i,t-1}} = \frac{\frac{RPS_{i,t}}{RPS_{i,0}}}{CPI_{i,t}} \times \frac{CPI_{i,t-1}}{\frac{RPS_{i,t-1}}{RPS_{i,0}}} = \frac{RPS_{i,t}}{RPS_{i,t-1}} \times \frac{CPI_{i,t-1}}{CPI_{i,t}} \tag{3}$$

I compute two series of the implied quality index. The first series is to apply equation (3) for all periods (implicit quality index: IQI hereafter). The second series is to apply equation (3) only for the period with the official announcement of the price substitution to the item (official quality index: OQI hereafter). The differences in the two series indicate implicit and undisclosed quality adjustment for discontinued individual price observations, not common to all observations for the item. Such effort is critical in maintaining the precision of the CPI at a high level, but, at the same time, it makes it difficult to make use of the time series information in the micro data for the RPS.

3.2 Simulation Results

Figure 2 provides the results for price substitution simulation by plotting the IQI and the OQI in December 2019 (the end of the sample period). The IQI and the OQI are log-transformed indexes for December 2019 by normalizing with 2015 at one, thus indicating the cumulative effects of quality adjustment at the price substitutions from January 2015 to December 2019.

The figure indicates three points: First, many observations are placed along with the 45-degree line, suggesting that many items show almost identical quality adjustment in both official and implicit price substitutions. Second, many other observations are located along the vertical axis, implying that implicit quality adjustments are made at individual price observations without the official announcement of price substitutions. Third, other observations that deviate from the vertical axis and 45-degree line suggest that both official and implicit substitutions are made, but quality improvement and deterioration are diverse across the items.

Figure 2: Results for Price Substitution Simulation

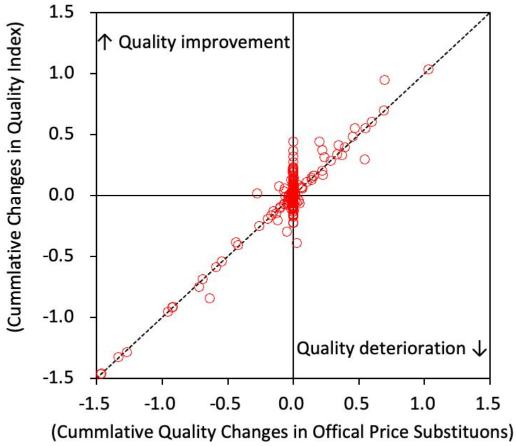


Figure 3 plots the data for horizontal and vertical axes in Figure 2 as frequency distributions in separate panels. The left panel of the figure for OQI shows a spike at exactly zero with more than 50 percent of items. The right panel for IQI also shows spikes around zero, but no observations exist at exactly zero, and two peaks exist for the intervals from -0.1 to zero and from zero to 0.1. That suggests that price substitutions at individual price observations are made at all the items, even though no official price substitutions at an item level are made for more than half of the items. Such adjustments in an individual price observation are critically important in maintaining the precision of price information in practice. That, however, makes it very difficult to employ time-series information in micro data for the RPS since it produces many discontinuities in individual price observations.

Figure 3: Frequency Distributions for Official and Observed Price Substitutions

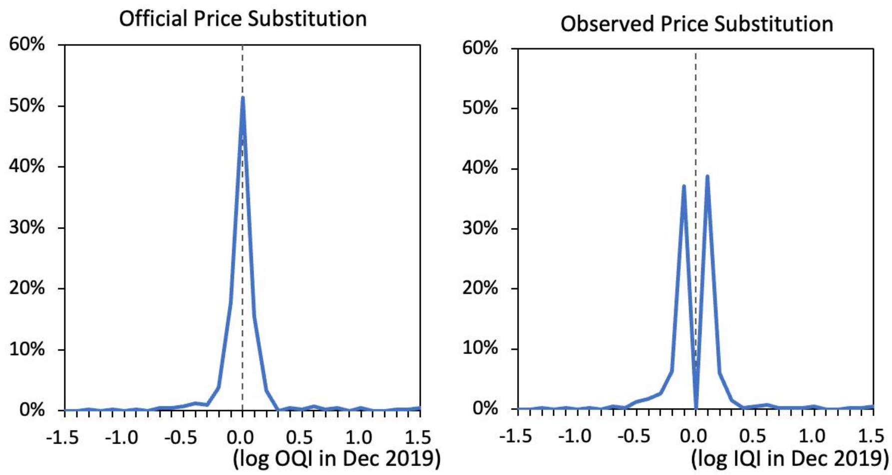
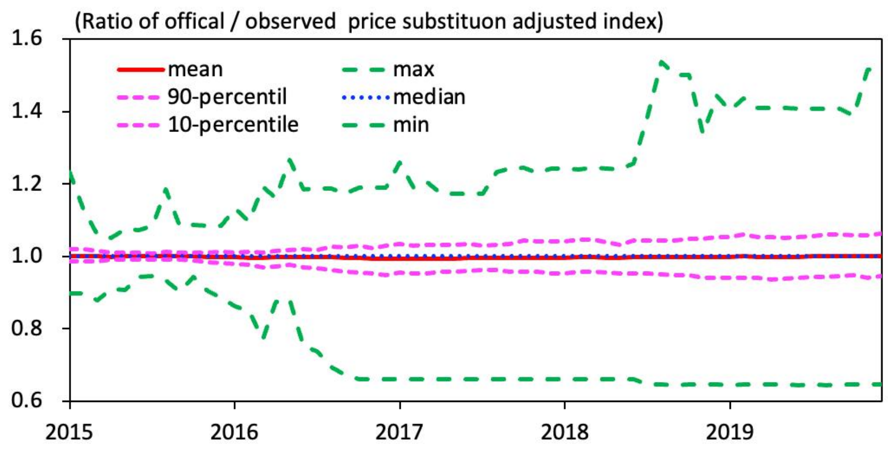


Figure 4 plots the distribution of the ratio of the IQI to the OQI with mean, median, max, min, 10-percentile, and 90-percentile values over the whole sample period from January 2015 to December 2019. The figure shows that the intervals between 10- an 90-percentiles are relatively tight, but the intervals become wider for the max and min values. That observation is consistent with Figure 3, suggesting the existence of some outlier items with significant differences between official and implicit price substitutions.

Figure 4: Cumulative Effects of Implicit Price Substitutions



Given the price substitution simulation results with discontinuous price observations, I decided to focus solely on cross-sectional information on price observations for each item, available from the micro data for the RPS.

4 Elementary Aggregation Bias

4.1 Elementary aggregation formula

The following three formulas are mainly used as an elementary aggregation formula, which aggregates individual price observations at the lowest level to the item price index.

$$I^C_{i,t} = \frac{1}{n}\sum_{j=1}^{n}\frac{p_{i,j,t}}{p_{i,j,0}}, \tag{4}$$

$$I^D_{i,t} = \frac{\frac{1}{n}\sum_{j=1}^{n} p_{i,j,t}}{\frac{1}{n}\sum_{j=1}^{n} p_{i,j,0}}, \tag{5}$$

$$I^J_{i,t} = \prod_{j=1}^{n}\left(\frac{p_{i,j,t}}{p_{i,j,0}}\right)^{\frac{1}{n}} = \frac{\prod_{j=1}^{n}(p_{i,j,t})^{\frac{1}{n}}}{\prod_{j=1}^{n}(p_{i,j,0})^{\frac{1}{n}}}, \tag{6}$$

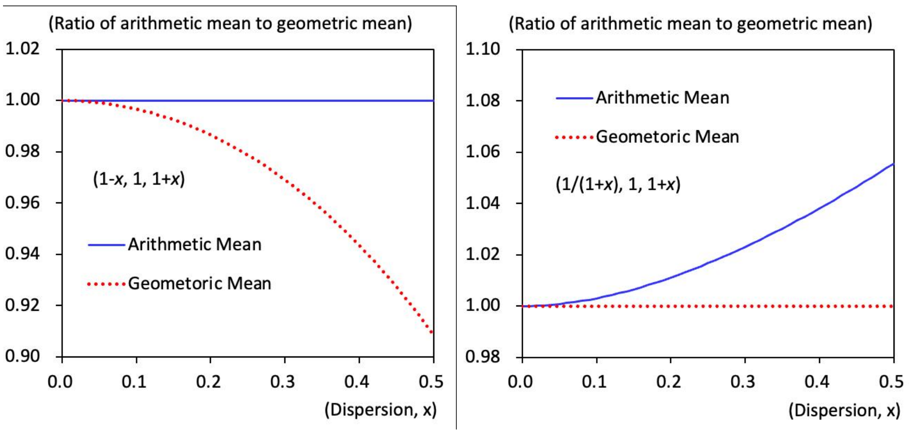
where $p_{i,j,t}$ denotes the price of j-th price observation in i-th item at time t, n is the number of price observations, and $t = 0$ is the base period.

Equation (4) is called the Carli index formula of the arithmetic mean of price relatives at the current period and the base period. Equation (5) is the Dutot index formula of the ratio of the arithmetic mean of prices at the current and base periods. Equation (6) is the Jevons index formula of the geometric mean of price relatives, comparable to equation (4), as well as the ratio of the geometric mean of prices at the current period and base period, comparable to equation (5).[6]

With very homogeneous price observations, all the elementary aggregation formulas produce very close results, and elementary aggregation bias, or narrowly defined lower-level substitution bias, is not a major issue. However, once price observations become heterogeneous and diverse, the elementary aggregation formula matters a lot.

That point becomes clearer as price dispersion becomes wide, as shown in two types of simple mean-preserving-spread exercises. First, the left-hand side panel in Figure 5 carries out a standard mean-preserving-spread exercise by computing arithmetic and geometric means for three observations $(1-x, 1, 1+x)$ by increasing x from zero to 0.5. In this exercise, the arithmetic mean remains unchanged while the geometric mean continues to decline as x becomes larger. Second, the right-hand side panel conducts a mean-preserving-spread exercise in terms of the geometric mean by computing both the arithmetic and geometric means for $(1/(1+x), 1, 1, x)$. In this case, the geometric mean remains unchanged, while the arithmetic mean continues to increase.

Figure 5: Difference between Arithmetic and Geometric Means



Note: Mean-preserving spreads exercise using data for (1-x, 1, 1+x) and (1/(1+x), 1, 1+x), respectively .

Two simple exercises shown above reveal that dispersion of price observations becomes larger, the arithmetic mean exceeds the geometric mean at an accelerated pace. That is a prac-

[6] International Monetary Fund (2020), which summarizes the international recommendations on the CPI compilation methods and practices, the Jevons index formula is preferred in most circumstances when weights are not available. The Dutot index formula can be employed only for homogeneous elementary aggregates since the price relatives are implicitly weighted by the price level in the reference period. However, Diewert (2021) strongly recommends using the Jevons index formula in any circumstances from the theoretical perspective.

As a more general formulation, the CES aggregator is an alternative index formula, which encompasses both the Dutot and Jevons index formulas, equations (5) and (6), with the elasticity of substitution σ of zero and one, respectively. Shapiro and Wilcox (1997) for the details of the CES index formula. Shiratsuka (1999) applies the CES index formula to the upper-level aggregation for the Japanese CPI as an alternative real-time index formula.

$$I_{i,t}^{CES} = \left[\sum_{j=1}^{n} \left(\frac{p_{i,j,t}}{p_{i,j,0}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$$

tical thing that the effects of elementary aggregation formula on measurement errors in the CPI. In that sense, the Japanese CPI is robust against the elementary aggregation bias since it employs the one-specification for one-item policy to keep price observations for each item homogeneous as possible.

4.2 Heterogeneity in price observations

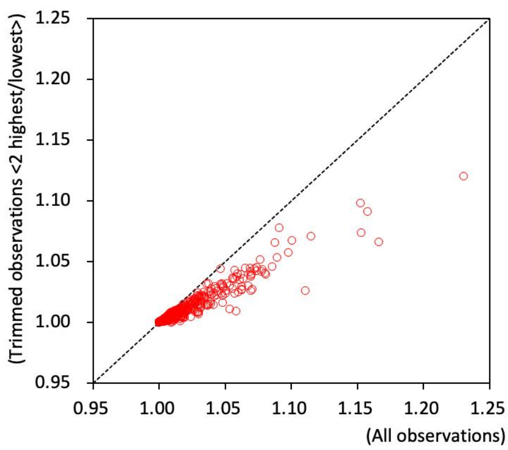
To examine the price heterogeneity at an item level, I introduce the ratio for arithmetic and geometric means (hereafter AG-ratio) for item i at period t, $AGR_{i,t}$, given by

$$AGR_{i,t} = \frac{\frac{1}{n}\sum_{j=1}^{n} p_{i,j,t}}{\prod_{j=1}^{n}(p_{i,j,t})^{\frac{1}{n}}}. \tag{7}$$

If all prices surveyed are identical, then AG-ratio will be one since the arithmetic mean and the geometric mean produce the same mean value. However, as price observations become diverse, the arithmetic mean constantly overstates the geometric mean, making the AG-ratio larger.

Figure 6 shows a scatter plot for the AG-ratio and the coefficient of variation (CV) for all common items in the CPI and the RPI by dividing them into two groups: mid-month survey items and every ten-day survey items. This figure clearly shows a positive correlation between the AG-ratio and CV. Focusing on the mid-month survey items, the relationship between the AG-ratio and CV seems nonlinear, and the AG-ratio increases at an accelerated pace as CV increases.

Figure 6: Ratio of Arithmetic-Geometric Means



The nonlinear relationship between the AG-ratio and CV becomes evident with panel regression results in Table 4. The AG-ratio is regressed on CV and its quadratic term by applying three panel data estimation procedures: pooled, fixed-effect, and between estimations.

$$AGR_{i,t} = Const. + \alpha CV_{i,t} + \beta CV_{i,t}^2 + \epsilon_{i,t},$$

where $CV_{i,t}$ denotes the coefficient of variation for item i at period t.

Looking at the table, pooled and between estimates are closer than fixed-effect estimates, suggesting that the between variation across items is more important than the within variation over time.

Table 4: Panel Regression on AG-Ratio

	Pooled			Fixed-effect			Between		
	Coeff.	S.E.		Coeff.	S.E.		Coeff.	S.E.	
CV	0.030	0.002	**	0.022	0.009	*	0.035	0.010	**
CV^2	0.406	0.006	**	0.427	0.023	**	0.394	0.024	**
$Constant$	0.998	0.000	**	0.998	0.001	*	0.997	0.001	**
obs.	23,825			23,825			418		
groups	—			418			—		
R^2	0.961			0.960			0.978		

4.3 Measurement bias from elementary aggregation formula

Next, I will compute annual inflation rates based on the arithmetic and geometric mean formulas, thereby estimating a lower-level substitution bias from their differences. It should be noted that this is a lower-level substitution bias in a narrow sense, examined by Boskin et al. (1996) for the U.S. CPI.

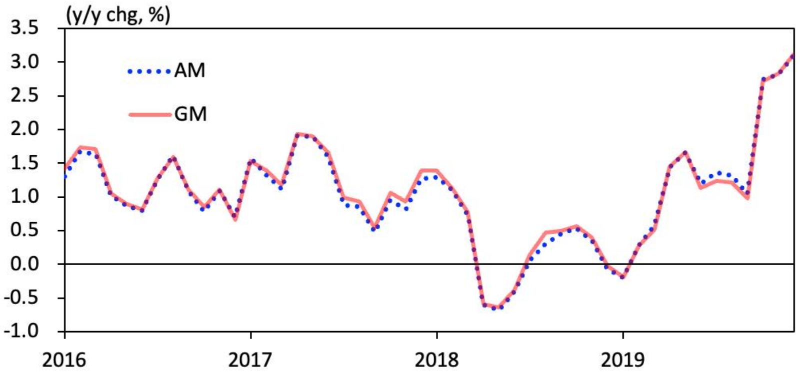
Figure 7 plots inflation rates for mid-month survey items in the upper panel and every ten-day survey items in lower panel. Differences are generally small, especially in the case of every ten-day items. Every ten-day survey covers volatile prices, mainly perishables influenced heavily by weather conditions but succeeds in controlling the price homogeneity.

As a result, the impact of lower-level substitution bias on the overall CPI inflation is also limited over time, as shown in Figure 8, with the maximum value of 0.08 percentage point in an absolute term. The figure also shows that the directions of the bias are unstable over time with both negative and positive values, even though the bias exhibits mostly negative values. The arithmetic mean constantly overstates the geometric mean, but their ratio varies over time, thereby pushing inflation rates upward and downward.[7]

[7] Very minor impacts from the elementary aggregation bias, as shown in this paper, do not necessarily imply that we do not need to reexamine the current practice of the Japanese CPI to apply the Dutot index formula to all the items in the elementary aggregation. As International Monetary Fund (2020) explains, the preferable elementary aggregation formula without weight information is the Jevons index formula, and the Dutot index formula is applicable only to the homogeneous elementary aggregate. Thus some items with highly heterogeneous and differentiated price observations must consider the elementary aggregation formula to the Jevons. It should be also noted that Diewert (2021) recommends not using arithmetic mean formula since it is not invariant to general changes in the units of measurement.

Figure 7: Difference in Inflation Rates between Arithmetic and Geometric Means

(a) Mid-month survey items



(b) Every 10-day survey items

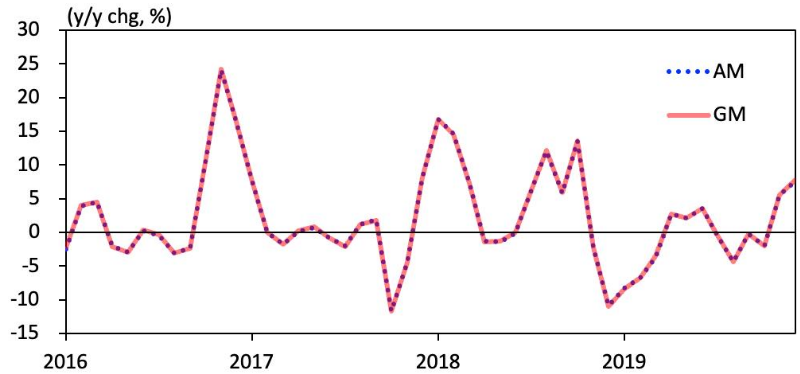
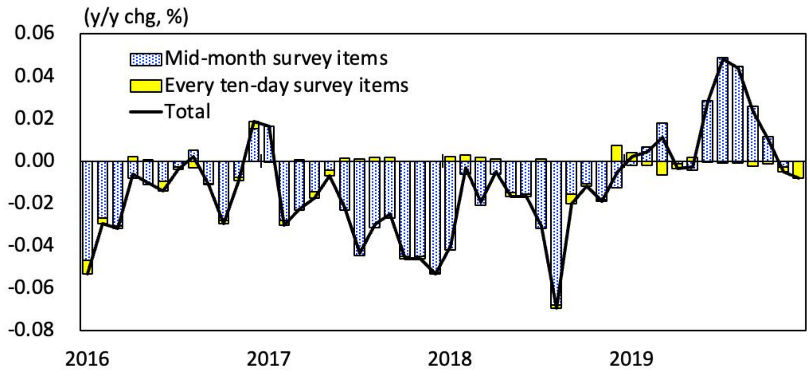


Figure 8: Elementary Aggregation Formula Bias



To sum up, lower-level substitution bias stemming from the elementary aggregation formula is relatively minor and produces only negligible impacts on the measurement errors in the Japanese CPI. However, the more fundamental problem exists regarding the lower-level substitution bias through the price representativeness due to the one-specification for one-item policy in the price survey.

5 Case Study on Heterogeneity of Price Observations

5.1 Framework for case study analysis

I carry out a case study analysis to examine the effects of elementary aggregation formulas using selected items based on the AG-ratio and the degree of detail in the specification. Table 5 summarizes selected items for case study analysis. Items from number one to three are stipulated with the detailed specification, while those from number four to nine are stipulated with the slightly loose specification. Items from ten to 12 are every ten-day survey items of perishables with different price volatility.

Table 5: Selected Items for Case Studies

No.	Items		Specification / Price volatility
(Mid-month Survey Items)			
1	Whisky	Detailed	Suntory Kaku Bin
2	Soy Sauce		Kikkoman or Yamasa (1-little bottle)
3	Washing Machine		Specific brand and product (undisclosed)
4	Butter	Loose	200-gram box (excluding unsalted butter)
5	Gasoline		Regular gasoline (1 little, excluding self-service station)
6	Cap		Standard product with logo (chemical fiber)
7	Women's Jeans		Standard product with logo (chemical fiber)
8	Pan		Saucepan (standard product, stainless steel, 18cm, 2.0-2.5 little)
9	Battery		AA alkaline battery (packed 4 batteries)
(Every Ten-day Survey Items)			
10	Sea Bream (Tai)	—	Low price volatility
11	Horse Mackerel (Aji)		High price volatility
12	Lettuce		

To examine the effects of price heterogeneity in each item, I compute the arithmetic and geometric means on a trimmed basis, given by

$$TRA(k)_{i,t} = \frac{1}{n-2k} \sum_{j=1+k}^{n-k} p_{i,j,t}, \tag{8}$$

$$TRG(k)_{i,t} = \prod_{j=1+k}^{n-k} (p_{i,j,t})^{1/(n-2k)}. \tag{9}$$

where k denotes the number of trimmed price observations in both upper and lower tails of price distribution at period t. k is determined based on the number of price observations in each item as follows.

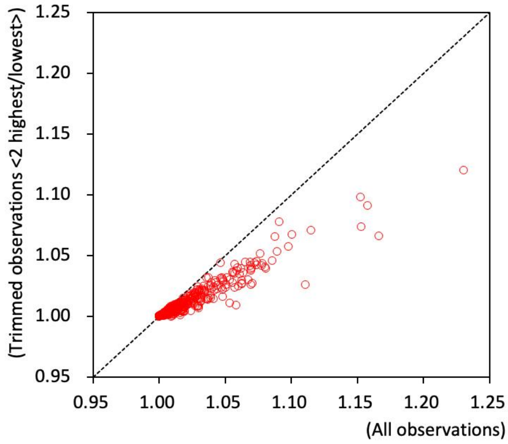
$$k = \begin{cases} 1 & (n = 12) \\ 2 & (n = 21). \\ 3 & (n = 42) \end{cases}$$

If price observations within the item become more homogeneous by trimming outlier observations, the arithmetic mean and the geometric mean on a trimmed basis show closer values, making the AG-ratio on a trimmed basis, given below, closer to one than the simple AG-ratio.[8]

$$AGR(k)_{i,t} = \frac{TRA(k)_{i,t}}{TRG(k)_{i,t}}. \tag{10}$$

Figure 9 exhibits a scatter plot for the AG-ratios for the simple and trimmed mean bases. The figure shows that all the observations are located below the 45-degree line, suggesting that price observations become more homogeneous in all items, but that the effectiveness of the trimmed mean procedure is limited in some outlier items.

Figure 9: AG-Ratio for simple and trimmed means



5.2 Case study analysis for mid-month survey items

I summarize the results for case study analysis for nine mid-month survey items in Figures 10. The figures for case study analysis includes the item CPI and five statistics: the arithmetic mean, the geometric mean, median, the trimmed arithmetic mean, and the trimmed geometric mean, which are normalized by the arithmetic mean for 2015 average being equal to 100 with a one standard error band around the arithmetic mean.

Whisky (Figure 10a)**:** The CPI for whisky is one of the most typical item for the price survey based on the one-specification for one-item policy. That specifies a single brand of *Suntory Kaku Bin*, one of Japan's most popular whiskey. As a result, arithmetic and geometric means are very close, suggesting fairly homogeneous price observations and resultant high precision of elementary aggregation. However, significant gaps between the simple and trimmed means exists from 2015 to 2017, indicating possible distortion from outlier price observation (high price observations in this case). Still, it is difficult to confirm whether the current CPI for whisky represents the overall price movements for the diversified whisky products in Japan. For example,

[8] If $k = 0$, equation 10 falls into equation 7 for the ratio of the simple arithmetic and geometric means.

the price hike in April 2016 reflects upward price revisions for Suntory's whisky products, and it is not clear whether it accurately reflects price trends for whisky products in general.

Soy Sauce (Figure 10b)**:** The CPI for soy sauces is also the typical item for the price survey based on the one-specification for one-item policy. It surveys two, not one, the most popular brands for soy sauce, *Kikkoman* and *Yamasa.* That makes price observations homogeneous, thereby achieving high precision of the elementary aggregation. It should be noted that small but constant deviations between the arithmetic and geometric means are observed, indicating the possible distortion from low price observations. And such deviations become smaller when using the the arithmetic and geometric means on a trimmed basis.

Washing Machine (Figure 10c)**:** The CPI for washing machines is another typical item for the price survey based on the one-specification for one-item policy, applied to goods and services with clear product cycles by periodic price substitutions. The figure shows strong seasonality in price fluctuations: prices jump in October and gradually decline afterward every year. The figure also shows that the CPI for washing machines remains generally flat, while other unit value indicators gradually increase over time, suggesting constant quality improvement. Even with such large price fluctuation with quality improvement, the figure suggests relatively high precision in the elementary aggregation. Still, it is difficult to confirm whether the current CPI for washing machines represents the overall price movements of the diversified washing machine market in Japan.

Butter (Figure 10d)**:** The CPI for butter has the relatively loose definition of the product specification, without clearly stipulating the brand, but seems successful in maintaining high precision in the elementary aggregation. However, it should be noted that some failures in the price survey, evidence as widening of the error bands for a couple of times, induce the distortions from outlier observations. In that respect, the arithmetic and geometric means on a trimmed basis show smoother movements at those times by limiting such outlier impacts.

Gasoline (Figure 10e)**:** The CPI for gasoline also successfully maintains high precision in elementary aggregation even with relatively large price fluctuations over time. Price differences across gas stations seem to be the outlet-specific factor, generally constant over time.

Cap (Figure 10a)**:** The CPI for caps achieves good precision in elementary aggregation even with relatively loose specification, saying "a standard product with logo (chemical fiber)." However, it shows a puzzling deviation between the CPI and the RPS in April 2017, indicating significant quality deterioration at that time.

Women's Jeans (Figure 10g)**:** The CPI for women's jeans shows a large AG-ratio and high coefficient of variation with a very wide range of price observations. It has a very loose specification, just saying "standard product for Blue denim jeans." The CPI contains the distortion from outlier prices through practical effort at individual price observations to some extent. However, it needs to review the price survey practice and examine the introduction of the geometric mean formula in the elementary aggregation.

Pan (Figure 10h)**:** The CPI for pans also shows a relatively large AG-ratio and high coefficient of variation with a wide range of price observations. It also shows a puzzling drop in the RPS in June 2018, while the CPI remains unchanged, suggesting significant quality deterioration at that time.

Battery (Figure 10i)**:** The CPI for batteries widens the price range after 2017, showing significant deterioration in the elementary aggregation precision. Such deterioration is attributable to the expansion of private label products in price observations, suggesting the difficulty in balancing the elementary aggregation precision and the price representativeness.

5.3 Case study analysis for every ten-day survey items

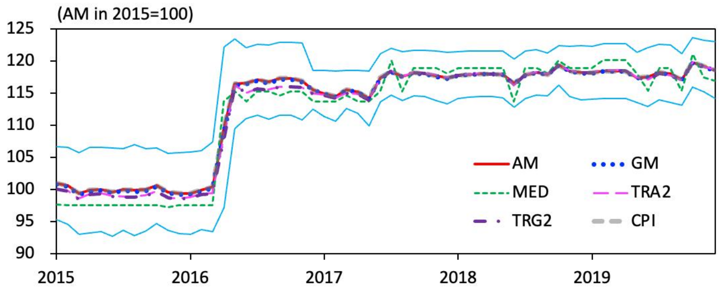
The case study analysis for every ten-day survey items include perishable foods with generally volatile price fluctuations due to changes in weather conditions.

I choose three items with slightly different volatility in price fluctuations: sea bream (*tai*) in Figure 11a**,** horse mackerel (*aji*) in Figure 11b, and lettuce in Figure 11c. Among the three items, sea bream shows relatively stable price fluctuation both in the short-term and longer-term due to increased aquaculture. Horse mackerel shows large short-term price volatility, but the range of price fluctuations is not so large in the longer term. Lettuce shows large swings in prices with seasonal fluctuations.

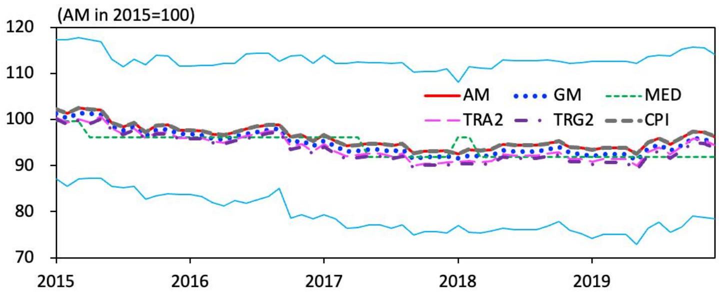
However, the three items with large price volatility show generally common movements for arithmetic and geometric means in both all observations and trimmed observations. That suggests that large price volatility does not necessarily produce distortionary effects from outlier price observations.

Figure 10: Case Study Analysis for Mid-Month Survey Items

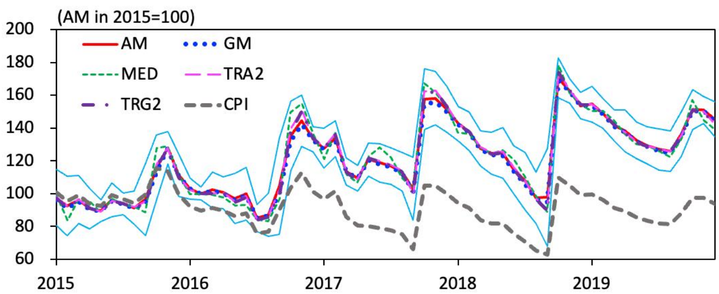
(a) Whisky (Item-1)



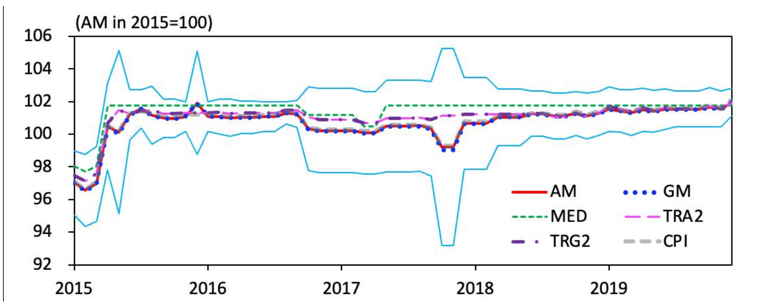
(b) Soy Sauce (Item-2)



(c) Washing Machine (Item-3)



(d) Butter (Item-4)



(e) Gasoline (Item-5)

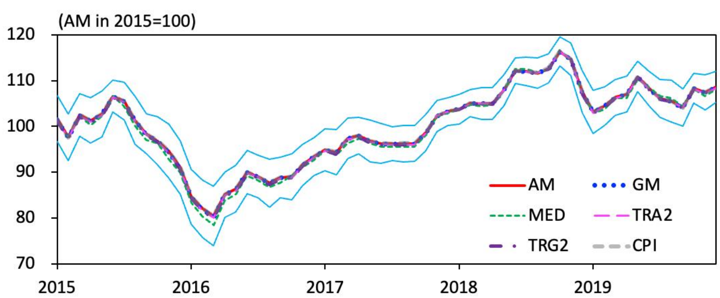
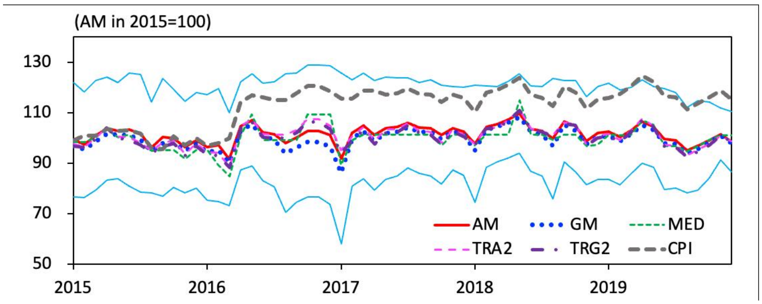
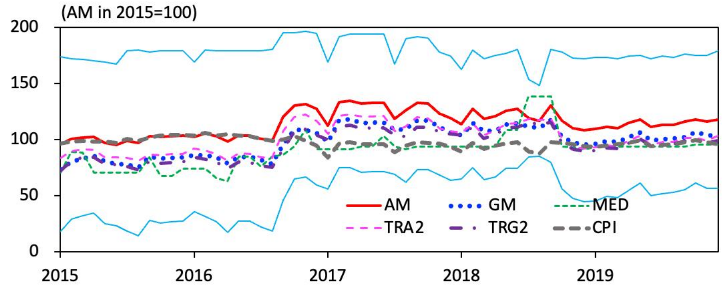


Figure 10: (Continued)

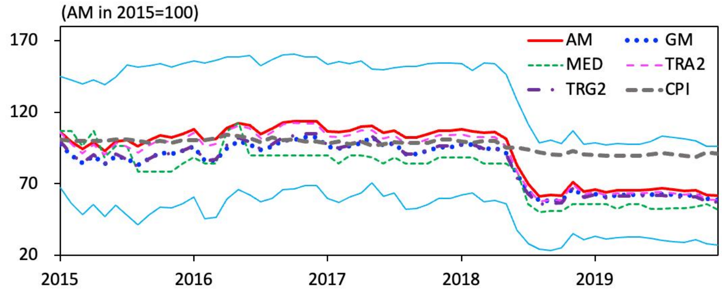
(f) Cap (Item-6)



(g) Women's Jeans (Item-7)



(h) Pan (Item-8)



(i) Battery (Item-9)

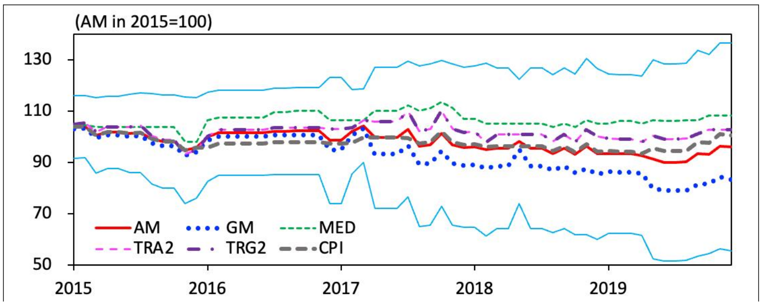
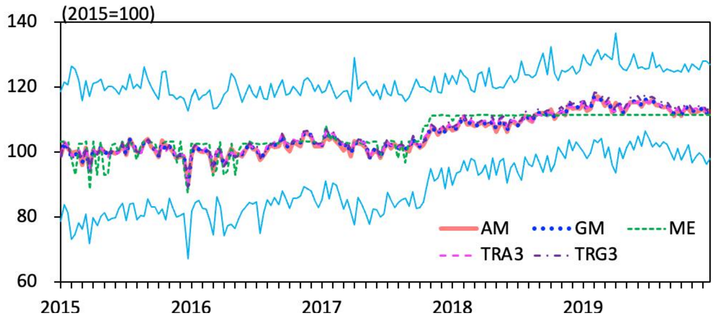
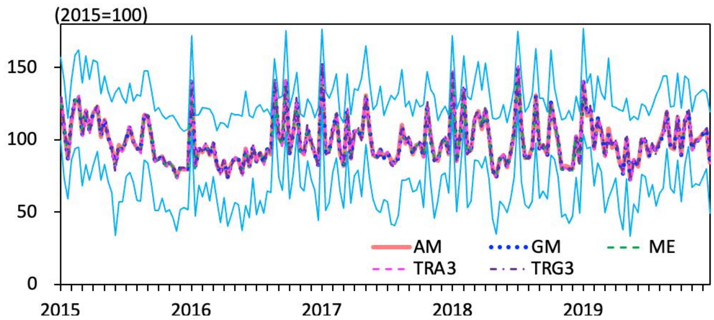


Figure 11: Case Study Analysis for Every-Ten-Day Survey Items

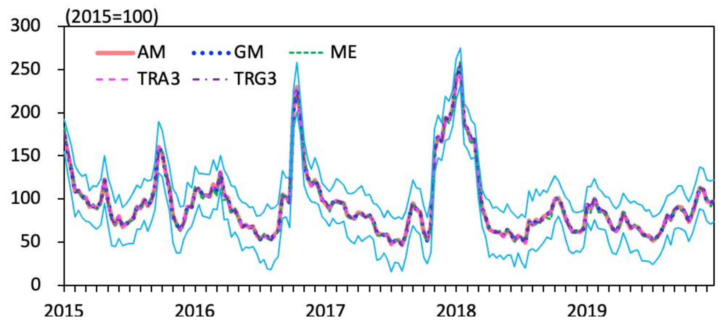
(a) Sea Bream: *Tai* (Item-10)



(b) Horse Mackerel: *Aji* (Item-11)



(c) Lettuce (Item-12)



6 Conclusions

In this paper, I explored the CPI measurement errors stemming from the lower-level substitution in the Japanese CPI and provided some quantitative evaluations.

With the first trial to employ the micro data for the RPS, I made it clear that the lower-level substitution bias in the Japanese CPI differs from that examined by Boskin et al. (1996) for the U.S. CPI. On the one hand, the Japanese CPI successfully constrains the elementary aggregation bias, or the narrowly-defined lower-level substitution bias, at a smaller margin than the U.S. CPI, reflecting the more homogeneous price observations within the items with the one-specification for one-item policy in the price survey. On the other hand, the Japanese CPI has problems of the broadly-defined lower-level substitution bias, which corresponds to the weakened price representativeness due to the one-specification for one-item policy. Some items with the loosely defined specification have features of highly heterogeneous and differentiated products within the item, resulting in the intensified price representative problems.

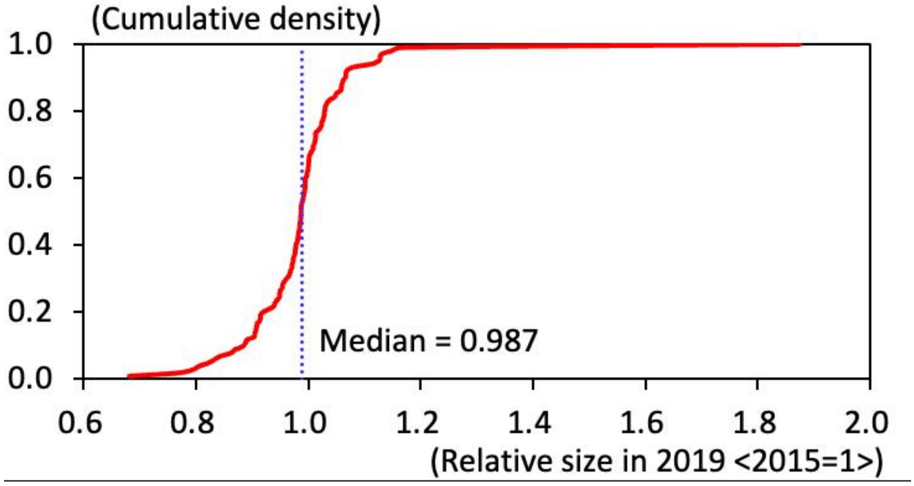
In addition, the current price survey method does not necessarily guarantee a stable price survey in practice. The micro data for the RPS reveals that it has many price discontinuities for temporary price survey failures due to the disappearance of the products surveyed at the outlets. Although the Statistics Bureau of Japan makes the best effort to link such discontinued price observations in compiling the CPI, it is undeniable that the Japanese CPI is vulnerable to outlier price observations stemming from the temporary price survey failure.

For further analysis on the price representativeness issues, it is necessary to employ alternative data sources, such as scanner data and web-scraping data. The Statistical Bureau of Japan has already employed such alternative data in a limited area, such as PCs, TV sets, and travel tour charges. It is deemed effective to widen the application of such alternative data step-by-step since the issues on the one-specification for one-item policy need to be examined on an item-by-item basis.

Concerning the price representativeness issues, another important point is how to deal with the downsizing of the products, reflecting the decrease in the household member size. Figure 12 plots the relative size of the items surveyed in 2019, normalized by the average size in 2015. The figure shows that the median size declines to 0.987, about 12 percent over four years. It is often argued that product prices typically remain unchanged with product downsizing, resulting in hidden price hikes.[9] However, it is not necessarily the case if household preference shifts toward products with smaller unit sizes. For example, single-person households prefer small-sized food staffs considering the risk of leftovers. In this case, higher unit prices for small-sized products reflect convenience premia compact and reasonable sizes. That is another open question that requires employing alternative data for a rigorous empirical analysis.

[9] Imai and Watanabe (2014) argue that the Japanese CPI is downwardly biased since the quality downgrade associated with product downsizing is not appropriately adjusted, based on their analysis using daily scanner data for supermarkets from 2000 to 2012. Abe, Inakura and Tonogi (2017) also point out that product turnover effects produce the significant price gap between new and old goods, using weekly scanner data covering general merchandise stores, supermarkets, drug stores and convenience stores from 2006 to 2016.

Figure 12: Distribution of Relative Size in 2019



Despite the efforts of statistical agencies in constructing accurate statistics, measurement errors are unavoidable to some extent in the CPI. Thus, the most important point is to ascertain whether they are small enough to be safely ignored in practice or serious enough to mislead users. The debate about the accuracy of the CPI should investigate the sources of measurement errors and how they affect accuracy. Making improvements in the CPI is difficult and requires a significant commitment of resources since each item needs to be examined separately.

References

Abe, Naohito, Noriko Inakura, and Akiyuki Tonogi, "Effects of the Entry and Exit of Products on Price Indexes," RCESR Discussion Paper Series DP17-2, Research Center for Economic and Social Risks, Institute of Economic Research, Hitotsubashi University March 2017.

Abraham, Katharine G., John S. Greenlees, and Brent R. Moulton, "Working to Improve the Consumer Price Index," *The Journal of Economic Perspectives*, 1998, *12* (1), 27–36.

Ariga, Kenn and Kenji Matsui, "Mismeasurement of the CPI," in "Structural Impediments to Growth in Japan" NBER Chapters, National Bureau of Economic Research, Inc, May 2003, pp. 89–154.

Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale Jorgenson, "Toward a More Accurate Measure of the Cost of Living," 1996.

__ , __ , __ , __ , and Dale W. Jorgenson, "Consumer Prices, the Consumer Price Index, and the Cost of Living," *The Journal of Economic Perspectives*, 1998, *12* (1), 3–26.

Broda, Christian and David E. Weinstein, "Defining Price Stability in Japan: A View from America," *Monetary and Economic Studies*, 2007, *25* (S-1), 169–189.

Deaton, Angus, "Getting Prices Right: What Should be Done?," *The Journal of Economic Perspectives*, 1998, *12* (1), 37–46.

Diewert, Erwein, "Elementary Indexes," 2021. mimeo.

Diewert, W. Erwin, "Index Number Issues in the Consumer Price Index," *The Journal of Economic Perspectives*, 1998, *12* (1), 47–58.

General Accounting Office, "Consumer Price Index: Update of Boskin Commission's Estimate of Bias," Technical Report, US General Accounting Office, 2000.

Greenlees, John S., "The BLS Response to the Boskin Commission Report," *International Productivity Monitor*, Spring 2006, *12*, 23–41.

Imai, Satoshi and Tsutomu Watanabe, "Product Downsizing and Hidden Price Increases: Evidence from Japan's Deflationary Period," *Asian Economic Policy Review*, January 2014, *9* (1), 69–89.

International Monetary Fund, *Consumer Price Index Manual: Concepts and Methods,* 2020.

Lebow, David E. and Jeremy B. Rudd, "Measurement Error in the Consumer Price Index: Where Do We Stand?," *Journal of Economic Literature*, 2003, *41* (1), 159–201.

Nordhaus, William D., "Quality Change in Price Indexes," *The Journal of Economic Perspectives*, 1998, *12* (1), 59–68.

Pollak, Robert A., "The Consumer Price Index: A Research Agenda and Three Proposals," *The Journal of Economic Perspectives*, 1998, *12* (1), 69–78.

Shapiro, Matthew D. and David W. Wilcox, "Alternative strategies for aggregating prices in the CPI," *Federal Bank of St. Louis Review*, 1997, pp. 113–125.

Shiratsuka, Shigenori, "Measurement Errors in the Japanese Consumer Price Index," *Monetary and Economic Studies*, 1999, *17* (3), 69–102.

__ , "Comment," *Monetary and Ecnomic Studies*, 2007, *25* (S-1), 195–204.