IMES DISCUSSION PAPER SERIES

Broad Perspective Review

Broad-Perspective Review Series

Measurement Errors in the Consumer Price Index: Perspectives on Numerical Targets for Price Stability in Major Economies

Satoshi Kobayashi, Takeshi Shinohara, Shigenori Shiratsuka, Nao Sudo, and Itofumi Takeuchi

Discussion Paper No. 2024-E-16

IMES

INSTITUTE FOR MONETARY AND ECONOMIC STUDIES

BANK OF JAPAN

2-1-1 NIHONBASHI-HONGOKUCHO

CHUO-KU, TOKYO 103-8660

JAPAN

IMES Discussion Paper Series 2024-E-16

November 2024

Measurement Errors in the Consumer Price Index: Perspectives on Numerical Targets for Price Stability in Major Economies

Satoshi Kobayashi,*1 Takeshi Shinohara,*2 Shigenori Shiratsuka,*3 Nao Sudo,*4 and Itofumi Takeuchi*5

Abstract

The consumer price index (CPI) is widely used by major central banks in the definition of a numerical target for price stability. In this context, the measurement errors, which became a significant concern following the U.S. Boskin Report in 1996, are still occasionally cited as one of the justifications for setting a positive inflation rate target. This paper presents a review of the current status of discussions and studies on measurement errors in major economies, including Japan, with a particular focus on those published since the Boskin Report. In light of the accumulated research to date, the implementation of solutions to enhance index accuracy in each country, and changes in economic structure, the magnitude of measurement errors that were pointed out in the report appears to have decreased overall, although there are differences across countries and types of errors. Nevertheless, certain areas, particularly those pertaining to service prices, continue to present challenges, and in consideration of the potential impact of changing economic structures, such as e-commerce, measurement errors may fluctuate significantly and remain unresolved in the transition towards a future increase in service consumption, a more digitalized economy, and an aging population. Given these issues, measurement errors could remain a valid basis for setting numerical targets for the CPI.

Keywords: Consumer Price Index; CPI; Measurement errors; Upper Bias; Quality Adjustment; Service Price

JEL classification: C43, E31

*1 Director and Senior Economist, Institute for Monetary and Economic Studies (IMES), Bank of Japan (BOJ) (E-mail: satoshi.kobayashi@boj.or.jp)

*2 Deputy Director and Economist, International Department, BOJ (takeshi.shinohara@boj.or.jp)

*3 Professor, Faculty of Economics, Keio University (E-mail: shigenori.shiratsuka@keio.jp)

*4 Deputy Director-General, Financial System and Bank Examination Department, BOJ (E-mail: nao.sudou@boj.or.jp)

*5 Economist, IMES, BOJ (E-mail: itofumi.takeuchi@boj.or.jp)

The authors would like to thank Yukinobu Kitamura, Jouchi Nakajima, Naoki Wakamori, Sohei Kaihatsu, Harumasa Shirakawa, Kazuhiro Hiraki, and the staff of the Bank of Japan. The views expressed in this paper are those of the authors and do not necessarily reflect the official views of the Bank of Japan.

1. Introduction

The accurate measurement of prices is of significant importance not only for economic analysis and compilation of statistics, but also for the broader field of economic activity. Price indices, which have been studied and calculated over many years to capture prices, can be likened to the "thermometer of the economy" and are used to evaluate the economic situation over time, as well as basic statistics in the System of National Accounts (SNA), etc. In some countries, the consumer price index (CPI) is employed as a price index when adjusting fiscal expenditure, as well as a reference index for price fixing in private transactions.

In principle, measurement errors (or biases) are defined as the difference between the actual statistics, CPI, and the price level measured based on the cost of living index (COLI).[1] There are many sources of measurement errors. For instance, the nature of the formulas used to aggregate the prices of individual goods and services, the limitations of statistical practice in capturing daily changes in consumer behavior and commercial distribution, and the emergence of new products and services can all lead to measurement errors.

The type and magnitude of measurement errors also have important implications in the field of monetary policy. To illustrate, if a central bank implements a policy aimed at price stability with the CPI as the policy target, and there are permanent measurement errors in the CPI, it might be natural to incorporate the size of the measurement errors or change in the measurement errors into the level of the numerical target for price stability itself. Indeed, in many countries, including Japan, the European Central Bank (ECB), Israel, Australia, Canada, Switzerland, and the Czech Republic, the CPI measurement errors are explicitly used as one of the reasons for a numerical target for price stability.[2]

One of the reasons for the increased attention paid to the measurement errors of the CPI was the publication of the report "Toward A More Accurate Measure of the Cost of

1 In addition to this approach, the measurement errors could be measured by comparison with a price index based on other criteria and axioms. The sources of measurement errors are generally the same regardless of the approach taken, and the steps taken to improve the CPI so that it captures consumer spending trends in a timely and appropriate manner are the same. Therefore, this paper follows the former COLI approach, and surveys the discussions, based on major previous studies. See Appendix 1 and Appendix 2 for details on the price index theory based on the axiomatic approach, the COLI concept, and the superlative index.

2 Academically, there are multiple factors on which the numerical target for the CPI is based. For more information, see Bank of Japan (2024).

Living" published by the Advisory Commission to Study the Consumer Price Index, initiated by the U.S. Senate Committee on Finance, known as the "Boskin Report" (Boskin *et al*., 1996). The report employed a quantitative methodology to evaluate the various factors contributing to measurement errors in the U.S. CPI, and the estimated measurement errors (commonly known as "upward bias") at the time was found to be approximately +1.1 percentage points. The Boskin Report was a pioneering study used by many countries, including Japan, in the analysis of CPI measurement errors.

This paper presents a survey of discussions and studies published since the Boskin Report on the concept, type and magnitude of measurement errors in the CPI, focusing on those related to Japan, the U.S. and Europe. Given the nature of the subject matter of measurement errors in the CPI, a large proportion of the research results are also from the statistics-producing authorities themselves. The concept of measurement errors has remained consistent over time and across countries. However, the magnitude of measurement errors are contingent upon the period under consideration, the economic structure of the country in question, and the statistical production methods employed. Consequently, there is some variation in the degree of measurement errors across time periods and countries.

The following are some preliminary conclusions. First, an overview of the trends since the release of the Boskin Report indicates that the total value of measurement errors, which was estimated to be +1.1 percentage points at that time, has been reduced overall in the U.S., Japan, and Europe. This reduction is the result of the solutions proposed in the report, although there are regional differences and differences in type of measurement errors. The reduction is due to the fact that, in each country, the authorities compiling the statistics have promoted solutions aimed at improving index accuracy in various areas. These include the selection of price survey targets, the method used to aggregate price indices for individual brands/products and items[3] into upper-level indices, and the frequency of updating the weights used in aggregation. Furthermore, the precipitous decline in the price of information and communication technology (ICT) equipment, which was conspicuous around the time of the report's release, has since abated. Such a

3 In this paper, the most detailed level of prices comprising price indices is defined as "brand/product" or "specification." The smallest published category of goods and services in which each specification is included is defined as "item," and the price index based on aggregate item indices is "overall" index. In the case of the item "chocolate," there are many types of chocolate of generally the same quality in the market, and the price index for each brand/product sold by each firm corresponds to the brand/product price index, and the price index for each brand/product is aggregated into an upper-level price index for the entire chocolate product, which is the item price index of "chocolate."

shift in the economic environment could be an influential factor in reducing measurement errors.

However, it does not mean that measurement errors have been eliminated and will not expand again in the future. Although the existence of measurement errors is acknowledged, there are certain item prices, primarily service prices, where measurement errors are expected to persist due to constraints such as the absence of a theoretically established measurement method and practical limitations such as the challenge of obtaining the requisite sample size when estimating measurement errors. Furthermore, the advent of digitalization has led to diversification and complexity in the ways in which firms provide goods and services and consumers pay for them. However, academic research on the implications of these changes in commercial distribution, particularly in terms of the types and magnitude of measurement errors, is not yet sufficient. Given these issues, it is not possible to rule out the possibility that measurement errors may fluctuate and expand in the future as a consequence of theoretical and practical advances, in addition to further changes in economic structure.

The following provides an overview of this paper. Section 2 describes the definition and types of measurement errors. Section 3 provides an outline of the Boskin Report and the research findings since the publication of the report. Section 4 summarizes the solutions introduced by national statistical authorities to improve the accuracy of the price index and changes in the economic structure after the publication of the report. Section 5 presents related research on remaining and new challenges. Section 6 concludes.

Technical aspects are mentioned in the appendices. In Appendix 1 and Appendix 2, we refer to price indexes based on axiological and econometric approaches. The axiomatic approach and economic approach derive the price index using the statistical properties and economic theory, respectively, as a starting point. This paper is primarily based on the economic approach. Appendix 3 describes the functional form of the major price indexes. Appendix 4 provides basic information on price indexes for each country, and Appendix 5 summarizes major efforts to improve index accuracy. The appendices provide a comprehensive summary of detailed information that could not be included in the main text. They should be read as appropriate. Appendix 6 presents estimates of the upper-level substitution bias for each country. Appendix 7 discusses the impact on price surveys of the increasing diversity and complexity of pricing.

2. Ideal Price Index and Measurement Errors

(1) What Prices Should Be Measured?

The benchmark in estimating measurement errors is COLI. COLI is calculated by determining a utility level of a certain magnitude, selecting the combination of goods and services that households need to achieve this magnitude, and then adding up the amount spent on goods and services corresponding to this combination[4] (more precisely, it refers to the combination that can achieve a certain level of utility, and for which the total sum of expenditures is minimized. See Appendix 2 for details). By calculating the fluctuations in this value over time, it is possible to identify the price changes that are necessary to maintain a certain level of utility.

Nevertheless, calculating and publishing COLI as a statistic on a regular basis is not straightforward from a practical standpoint. As prices of goods and services fluctuate on a daily basis, the "combination of goods and services to maintain a certain level of utility" is also considered to change on a daily basis, reflecting shifts in demand from higher-priced items to lower-priced items. Consequently, in order to calculate COLI, it is essential to have information on the manner in which individual expenditure weights vary, as well as individual prices. Depending on the shape of the assumed utility function, parameter values for the function are also needed.

It is also essential to recognize the need to capture the portion of observed price variation that excludes the quality variation portion. This is because even if the apparent price per unit of goods and services provided remains constant, some consumers' compensation per unit of utility fluctuates due to changes in the quality of the goods and services. In the context of the COLI concept, which is concerned with the capture of changes in expenditure when the level of utility remains constant, it is theoretically necessary to capture changes in compensation for goods and services that provide the same level of utility. In many cases, the amount of quality variation cannot be directly observed and must be estimated or assumed.

4 The total expenditures is the sum of value, which is equal to the quantity of goods and services purchased by consumers times their prices; theoretically, expenditure is the function of individual prices and household income. Thus, COLI is also the function of those variables. Assuming a homothetic utility function, changes in COLI are independent of income changes in each period and are the function of individual prices and expenditure weights only.

One potential solution to the practical difficulties is to use the "superlative index" (Diewert, 1976) as an approximation of COLI. The superlative index is an index calculated exclusively from the prices and expenditures of individual goods and services, and this price index has the advantage of approximating a price index that is "exact[5]" to COLI, even in the absence of information on parameters related to the utility function. However, calculating and publishing the superlative index is not easy in practice. For example, the calculation of the Törnqvist index, one of the superlative indices, necessitates the expenditure weights at the base period (the point of comparison against which price indices are measured) and at the comparison period[6] (see the functional form of the index in Appendix 3). Consequently, if this price index is to be calculated and published on a monthly basis, it is necessary to obtain and compile the monthly expenditure weights in real time.[7, 8]

5 "Exact" means that the index matches the COLI under certain conditions. An index that is exact is called an "exact index."

6 In this paper, in accordance with the terminology used in general price statistics, when calculating a price index for a good or service, the reference point in time is called the "base period" and the point in time when the price level needs to be known (in most cases, the current point in time) is called the "comparison period." The calculated price ratio between the base period and the comparison periods is called an "index" and is expressed as "specification/brand/product index" when the aggregation level is for specification/brand/product and "item index" when the aggregation level is for items.

7 In general, it takes considerable time for not only price statistics, but also other statistics to be calculated and published. Surveyors and other staff review the answer written on the survey forms, filter out erroneous responses, enter the data, aggregate the data into upper-level indices, and publish the data. If expenditure weights are changed on a monthly basis, this work cycle must be synchronized with the release of expenditure weight statistics. Of course, if quarterly published statistics are used to calculate the weights, it may not be possible to update the weights on a monthly basis.

8 One of the reasons why the CPI is a cost of goods index (COGI) and is calculated using a fixed-weight base Laspeyres index in many countries might be that it is difficult to establish and maintain a monthly publication system even if the superlative index is adopted for the main price index. Nevertheless, reflecting growing public concern about measurement errors, some countries have published the superlative index or other indices (e.g., chained indices) that are considered to have smaller measurement errors than the fixed-weight base Laspeyres index as a reference series. By comparing such a reference series with the published main index, the fixed-weight base Laspeyres index, an environment is being created in which it is easy to observe (part of) the measurement errors timely.

Following the above discussions, in many countries, the CPI is positioned as a cost of goods index (COGI) rather than a COLI.[9] The COGI is a price index that captures how "the amount of expenditure required to consume a fixed combination (basket) of goods and services" changes over time, and in most cases, the fixed-weight Laspeyres index is used as the specific calculation method. In contrast to the COLI, the COGI uses fixed contents in its consumption basket, so it is only necessary to capture movements in individual prices for the calculation of the index. This results in a high frequency of publication without a significant lag between the time of the transaction and the subsequent publication. In this sense, COGI has practical advantages. However, there is no divergence between COGI and COLI in terms of measuring the change in price for a certain utility level rather than the price consumers observe, and even if COGI is used, it is necessary to take into account quality change. Consequently, even when the CPI is calculated as COGI, a quantitative evaluation of quality variation is conducted using various estimation and assumption methods, including the hedonic method described below.

9 In Japan, the Statistics Bureau of the Management and Coordination Agency (now the Statistics Bureau of the Ministry of Internal Affairs and Communications) stated, in 2000, that the purpose of the consumer price index is to measure the average change in prices of goods and services traded in the market and purchased by consumers. The cost of living index theory is an abstract hypothesis that assumes the existence of a utility function and mutually independent and rational consumption behavior. In the U.S., the Bureau of Labor Statistics (BLS), which prior to the release of the Boskin Report took the position that the CPI was COGI, now takes the COLI position ("we use a cost-of-living framework in making practical decisions about questions that arise in constructing the CPI"). However, in order to calculate the COLI exactly, it is necessary to capture the movements of other factors that affect utility, such as public goods and environmental goods, etc. Since it is technically difficult to capture such factors, the COLI is considered to be a conditional cost-of-living index in the U.S. The Harmonized Index of Consumer Price (HICP) in Europe is positioned as a COGI. However, it is recognized that conditional COLI, which excludes environmental goods, etc. that are difficult to measure, approximates COGI measured under ideal conditions where lower-level aggregation, sampling, and quality adjustment are similar (ECB, 2021).

(2) Definition of Measurement Errors and Types of Bias in Price Indexes

The measurement errors in the CPI are the difference between the price index published as a statistic (often COGI) and the price index that we really want to know (COLI). To summarize, measurement errors can be conceptualized as occurring along two distinct pathways. First, they arise from the CPI's inability to capture shifts in demand. Second, they emerge from the CPI's inability to accurately quantify the contribution of changes in quality to price changes. When the price of particular goods increases, households are expected to attempt to maintain their utility level by shifting their expenditures to other, cheaper goods. However, if the CPI fails to capture such demand shifts, it will capture a larger impact of goods whose prices have risen more than the price actually paid by consumers, resulting in measurement errors. Moreover, as previously stated, measurement errors will ensue in the event that alterations in quality cannot be accurately identified. Also, even identical goods and services may be priced differently at different points in time when they are sold (dynamic pricing), which can result in measurement errors when there is a discrepancy between the actual consumption behavior and the statistical samples and methods of the survey. The developments surrounding dynamic pricing are noted in Appendix 7.

The following paragraphs present four biases concerning measurement errors that have been frequently discussed in previous research: "upper-level substitution bias," "lower-level substitution bias," "quality change and new products bias," and "new outlet bias." The term "bias" as used in this paper is consistent with the usage of the word in previous studies on measurement errors in price indexes. The "bias" does not necessarily imply a specific direction of bias, but when measurement errors are biased in a specific direction, it will be described as "upward" or "downward" bias in this paper.

Upper-level substitution bias[10] is a form of bias that arises from the aggregation of item indexes, which occurs when the upper indexes (subgroup index, all-items indexes, etc.) are aggregated from the item indexes by weighted averaging. For example, if the price of margarine rises and the price of butter falls, demand may shift from margarine to butter. If the basket is fixed at the value prior to the change, the price of margarine is given greater weight than actual expenditures in the aggregation stage for the upper index. Consequently, the price level and price change rate increase. As the price level and change

10 In this paper, biases derived from index formulas (e.g., reset effects, drift phenomena, etc.) are also included in the upper-level substitution bias. See Appendix 6 for more information on the reset effect and drift phenomenon.

rate of margarine increase, the discrepancy between the aggregated upper-level price index and the true price index with respect to both level and change rate also increases in the positive direction. In other words, the failure to capture shifts to other goods and services as a result of price increases for certain goods and services introduces an upward bias in the price index.[11]

Lower-level substitution bias represents a substitution bias within the item index, occurring when the item index is calculated through the weighted and simple averaging of each price index or each price, which is the lowest level of aggregation.[12] This bias can occur when demand shifts to another brand or product with a lower price within the same item index. For example, in the case of the item "chocolate," if the market share of the cheaper chocolate becomes greater than that of the chocolate whose brand is surveyed in the CPI, an upward bias will occur unless the surveyed brand is switched to the cheaper one. In particular, if the range of qualities (volume, brand, function, etc.) of each price within the same item index is large, there may be differences in the price level (or index level) of each of the brands/products surveyed. In such cases, when aggregation methods such as simple arithmetic averages are used, greater weight is given to the movements of prices with higher price levels or index levels (Leifer, 1999). This bias is also expected to result in an upward bias in the CPI.

In order to ascertain changes in price over time, it is essential to adjust for change in the quality component of the price change. If the adjustment is not adequate, a **quality change bias** can result. With the exception of items such as ICT devices, for which product specifications are easy to find and vertical differentiation is a common practice, it is not easy to keep the quality of price-surveyed products fixed or to adjust for changes in quality when new products are introduced and old products are replaced. Although an upward bias is generally considered to occur for item indexes in which product quality is improved on a trend basis, it should be noted that the direction of quality change is not

11 Here, we consider normal goods. For normal goods, a price decline increases demand due to two factors: (1) an increase in real income due to lower prices (income effect) and (2) an increase in substitute demand due to lower relative prices (substitution effect). In the case of Giffen goods, (1) there is a decrease in demand due to an increase in real income, income effect, caused by lower prices (nature of inferior goods) and (2) there is an increase in substitute demand caused by lower relative prices (substitution effect), and as a result of the former (1) exceeding the latter (2), lower (higher) prices make lower (higher) demand.

12 There are several methods used to aggregate the prices of each brand/product. If the classification is by formula, there are arithmetic and geometric averages. If the classification is by the order of indexation, there are two methods: one is to calculate the average value of each index after indexing the surveyed brand/product prices, and the other is to index after averaging the price levels of surveyed brands/products.

always constant over time for some items, and an upward bias may not always be in effect.

New product bias is defined as one of the statistical errors associated with the continued surveying of outdated product prices that have lost their price representativeness, as well as with the omission of new forms of goods and services. Compared to older products, new products are often of higher quality for the price, which means that quality-adjusted prices of new ones are lower. Therefore, if new products cannot be included in the statistics, the index cannot reflect the "price level difference between old and new products" or "price reduction immediately after the launch of a new product," which can cause an upward bias.

New outlet bias is one of the measurement errors associated with changes in commercial distribution. This bias occurs when consumers shift to outlets and sales channels that sell and offer goods and services at lower prices and in higher quality sales formats, and when the CPI fails to capture these movements. As Shiratsuka (1999) points out, since the 1990s, Japan has seen a shift from individual stores in shopping malls to large supermarkets. While the quality of sales service (opening hours, convenience of location, etc.) is anticipated to vary by business type, if consumers perceive that the sales price in the new purchase channel is more affordable even after considering quality-related factors, they are likely to shift their purchasing channel and exhibit an upward bias, which is similar to the mechanism of other substitution biases.[13]

3. Historical and Recent Trends in Measurement Errors

A milestone in the study of CPI measurement errors was the publication of the Boskin Report. The Senate Finance Committee tasked the Advisory Commission to Study the Consumer Price Index, chaired by Professor Boskin of Stanford University, with examining the potential upward bias of the CPI and producing a report on the matter. In the U.S., social security spending was linked to the CPI (price indexing), and any measurement errors in the CPI could have a significant impact on the amount of fiscal spending. Since around the early 1990s, there has been a notable surge in the examination of measurement errors, including studies conducted by researchers at the Bureau of Labor Statistics (BLS), such as Reinsdorf, 1993; Darby, 1995; Diewert, 1995; Gordon, 1995; Jorgenson, 1995; Griliches, 1995. Also, then-Chairman of the Federal Reserve Board,

13 Gordon (2006) points out that, in general, the shift from full-price outlets to discount outlets is based on consumers comparing both price and service differences.

Alan Greenspan made remarks during a congressional hearing in January 1995, and he highlighted the existence of an upward bias in the CPI and its implications for the budget deficit. This context shaped the release of the Boskin Report (Kliesen, 1997).

The Boskin Report estimated that there was an upward bias of +1.1 percentage points in the overall U.S. CPI, which was calculated by summing the estimation results of four separate biases: upper-level substitution bias, lower-level substitution bias, quality change and new products bias, and new outlet bias (Chart 1). Based on estimates by the Congressional Budget Office (CBO), the report also states that if the +1.1 percentage point bias were to continue for the next 10 years, it would result in an increase of the 2006 budget deficit by the equivalent of $148 billion, due to higher social security and other costs.

Chart 1 Measurement Errors in the U.S. CPI Noted in the Boskin Report

(Annual rate, percentage points)

Substitution Bias	Upper-level Substitution Bias	0.15
	Lower-level Substitution Bias	0.25
Quality Change and New Product Bias		0.60
New Outlet Bias		0.10
Total		**1.10**

Source: Boskin *et al*. (1996).

Given the absence of an independent budget for the research and analysis in the preparation of the Boskin Report, the advisory commission did not adopt the approach of estimating the measurement errors. Instead, it aggregated the results of a diverse range of academic research that was available at the time, including studies published by the BLS and other public institutions. By reference findings from previous studies, the advisory commission quantified the magnitude of each type of bias.

First, in the report, the upper-level substitution bias was measured according to the analytical method proposed by Aizcorbe and Jackman (1993). In order to aggregate the indices into higher-level categories, Aizcorbe and Jackman (1993) calculated several indices based on the Laspeyres index with fixed weights used in the CPI, as well as the Fisher index and the Törnqvist index with variable weights, in the sense that the expenditure ratio in the comparison period was used. Quantitative comparisons were then made between them. It was also organized so that the effect on the price index of the

demand shift resulting from relative price changes of goods and services could be quantified by the deviation between the fixed-weight base Laspeyres index and the superlative index, the Törnqvist index. The Boskin Report also followed this approach, estimating the value of the bias at +0.15 percentage points, based on the latest estimates by the BLS (their internal data).

The lower-level substitution bias was based on the analysis of Shapiro and Wilcox (1996) (and previous studies: Moulton, 1993; Reinsdorf and Moulton, 1996; Moulton and Smedley, 1995). In the U.S. at the time, the BLS employed an arithmetic mean when aggregating from the individual price indexes to the item index. However, it was pointed out that the arithmetic mean does not capture the substitution effect, whereby higher-priced products are purchased in smaller quantities, thus creating an upward bias similar to the upper-level substitution bias. On the other hand, aggregation by geometric mean can capture a certain degree of substitution effect.[14] Therefore, while taking into account the +0.49 percentage points of bias reported by Moulton and Smedley (1995), which measured the discrepancy between the arithmetic and geometric means, and also taking into account subsequent changes in the lower-level aggregation methodology after this research, the Boskin Report assessed the value of the bias at +0.25 percentage points.

Regarding quality change and new product bias, the authors first divided the CPI into 27 categories and calculated the bias for each category using the quality change bias, etc. reported in previous studies. They then aggregated these biases for each category using expenditure weights, resulting in an overall value of +0.6 percentage points.[15] However, not all categories' errors were derived from previous academic research: in contrast to "goods" such as electrical appliances, clothing, and cars, for which previous studies were cited and quantitative evaluations were based on them, the number of items for which

14 Aggregation by geometric mean is equivalent to assuming an elastic utility function such that the nominal expenditure share does not change with changes in price, i.e., a function with a price elasticity of demand of 1 (1% increase in price decreases the quantity demanded by 1%).

15 Items for which measurement errors were found to exist are "appliances incl. electronic," "other utilities, incl. telephone," "prescription drugs," "food at home other than produce," "fresh fruits and vegetables," "shelter," "professional medical services," "hospital and related services," etc. Some items are considered to have zero measurement errors, such as "other private transportation" (automobile maintenance and repair costs, insurance fees, etc.), "health insurance," "housekeeping services," "cigarettes," and "fuels."

biases were calculated based on rough estimates were relatively high for "services", which may reflect the paucity of previous studies in this area.[16]

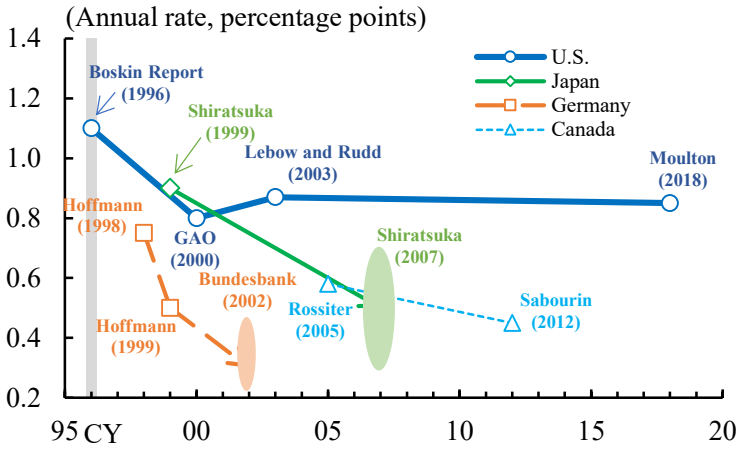
Finally, new outlet bias was calculated by the difference in the price of the same goods or services between purchasing outlets.[17] At that time, the U.S. CPI employed a sample rotation of one-fifth of the total sample on an annual basis, and Reinsdorf (1993) reported a 0.25 percentage point discrepancy in the survey prices of food products obtained from actual stores before and after the sample rotation. Based on this estimate and the study conducted by Lebow, Roberts, and Stockton (1994), which determined that approximately 40% of the overall CPI was affected by new outlet bias, the Boskin Report proceeded to undertake a quantitative assessment.

The Boskin Report was a milestone, and attempts were made to quantify the measurement errors of the CPI in countries other than the U.S., mainly from the late 1990s to early 2000s. In Chart 2, the results of the measurement error estimates by country are presented in a time-series format. This chart confirms that upward bias was reported in other countries as well, including Japan, as is clear from initial observations. For example, Shiratsuka (1999) employed the Boskin Report methodology to analyze Japan's 1995-base CPI, resulting in a reported bias of +0.9 percentage points. Furthermore, it can be confirmed that the magnitude of measurement errors is smaller in more recent years, with a particularly pronounced trend observed in Japan, Germany, and Canada. To illustrate, the measurement errors of the Canadian CPI decreased by 0.13 percentage points from the findings of Rossiter (2005) to those of Sabourin (2012).

16 For example, in "food away from home," they assumed that supermarket product assortments had improved living quality by 10% over the past 30 years and used the annualized value of that improvement. In "personal and educational expenses," it was assumed that personal financial services had improved living quality by 2% per year as a result of the spread of ATMs, etc. For "shelter," which accounts for a large weight, they applied the quality change bias of durable goods (home appliances) reported in Gordon (1990). In contrast, some biases were not based on analogy but on previous estimates, such as those for medical services based on the estimates of Cutler *et al*. (1996).

17 As mentioned above, when surveyed stores change, it is necessary to adjust differences in store facilities, etc., from the price level of the item as quality differences. For example, the size of parking lots and the number of tenants in suburban areas, and accessibility such as proximity to train stations and opening hours in urban areas, correspond to quality differences. If all price differences between stores for the same goods and services were due only to these quality differences in the stores' services, then there would be no new outlet bias. In this regard, Gordon (2006) suggested that, although there had been a shift from Sears and Kmart, which sell high-priced products, to Walmart and Target, which sell low-priced products, both of these were self-service stores, and Gordon (2006) mentioned that the difference in service quality between the two types of outlets in the U.S. had been marginal for this reason.

Chart 2 Measurement Errors by Country



Note: For previous studies where directional and upper range limits are indicated, these are shown shaded. GAO (2000) is the average of estimates by each member of the Advisory Commission to Study the Consumer Price Index.

Sources: Boskin *et al.* (1996); Bundesbank (2002); GAO (2000); Hoffmann (1998, 1999); Lebow and Rudd (2003); Moulton (2018); Rossiter (2005); Sabourin (2012); Shiratsuka (1999, 2007).

Chart 3 reorganizes these studies by factor. First, it can be pointed out that across countries, the quality change and new product bias is larger than other biases. For example, the Boskin Report and Shiratsuka (1999) demonstrated that while the overall measurement errors were +1.1 and +0.9 percentage points, respectively, quality change and new product bias accounted for the majority, +0.6 and +0.7 percentage points. Furthermore, this bias has been reduced in more recent studies. In a follow-up to the Boskin report, Lebow and Rudd (2003) observed a decline in the quality change and new product bias in the U.S., and estimated the value at 0.37 percentage points. Additionally, Shiratsuka (2007) documented a similar trend in Japan. Other biases were originally limited in magnitude, and some of them have declined further in recent years. However, the U.S. upper-level substitution bias has fluctuated up and down, with an expansion to some extent after the Boskin Report.

Chart 3 Previous Studies on Measurement Errors of CPI by Country

(Annual rate, percentage points)

	Japan		U.S.				Germany			France	Europe	Canada	
	Shira-tsuka (1999)	Shira-tsuka (2007)	Boskin Report (1996)	GAO (2000)	Lebow and Rudd (2003)	Moulton (2018)	Hoff-mann (1998)	Hoff-mann (1999)	Bundes-bank (2002)	Lequ-iller (1997)	Wynne (2005)	Rossiter (2005)	Sabourin (2012)
Upper-level substitution bias	0.00	→	0.15	0.10	0.30	0.25	0.10	0.10	—	0.05-0.10	—	0.15	0.22
Lower-level substitution bias	0.10	→	0.25	0.05	0.05	0.05	—	—	—	—	—		
Quality change and new product bias	0.70	↓	0.60	0.55	0.37	0.37	<0.60	<0.30	—	—	—	0.35	0.19
New outlet bias	0.10	↘	0.10	0.10	0.05	0.08	<0.10	<0.10	—	0.05-0.15	—	0.08	0.04
Total	**0.90**	↓	**1.10**	**0.80**	**0.87**	**0.85**	**0.75**	**0.5**	**<0.5**	**0.10-0.25**	**1.0-1.5**	**0.58 (-0.75)**	**0.45 (-0.60)**

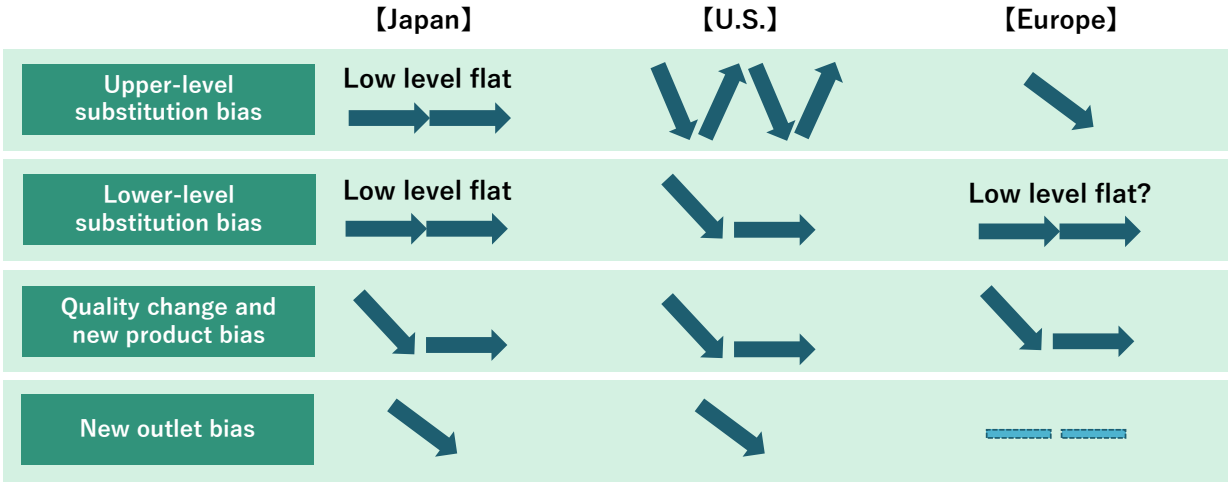
Note: Estimates by Shiratsuka (1999) are as of the 1995-base CPI. Arrows in Shiratsuka (2007) are from the original paper. Quality change and new product bias is indicated as "↓" because the quality adjustment method was expanded when the CPI base was revised. Although relatively smaller than its impact, the new outlet bias is also decreasing due to the inclusion of large suburban stores, "↘". Parentheses in Canadian columns are maximum estimates. Total value by GAO (2000) is an average of estimates by each member of the Advisory Commission to Study the Consumer Price Index, where the breakdowns are approximate calculations. Lebow and Rudd (2003) and Moulton (2018) report an upward +0.1 percentage point expenditure weight bias, so the sum of each item's value in this chart does not match the total value. In addition to these studies, there are studies focusing on Japan by Ariga and Matsui (2003) and Suga (2005), and on the U.S. by Broda and Weinstein (2007), which are not included in the table.

4. Prescriptions for Measurement Errors and the Recent Situation

Chart 4 summarizes the direction in Japan, the U.S., and Europe with respect to changes in each of the biases identified in the Boskin Report, based on an evaluation of previous studies. In addition to the effects of solutions to improve the accuracy of statistical indices implemented in various countries after the report, the discussion in Appendix 6 on upper-level substitution bias is also taken into account in this chart.[18]

First, the **upper-level substitution bias** was originally low for Japan and Europe and has since leveled off at a low level or appears to be trending downward. As for the U.S., it is seen fluctuating up and down at relatively high levels. The **lower-level substitution bias** is seen to remain low in both Japan and Europe. For the U.S., the bias seems to have decreased, reflecting solutions to improve statistics. **Quality change and new product bias** are expected to be on a decreasing trend. **New outlet bias** appears to have narrowed in Japan and the U.S.[19]

Chart 4 Direction of Measurement Errors



Note: "↘" means "declining trend," and "→" means "no change after a certain point." For example, in the case of "↘→", it means that the bias shrank to a certain point and then remains at that level. The U.S. upper-level substitution bias has fluctuated widely up and down.

Overall, improvements in statistical methods can be identified as the reason for the reduction in measurement errors. For example, in the U.S., following the release of the Boskin Report, there were active discussions between economists in academia and

18 In this chart, we evaluate only the quality change/new product bias and new outlet bias, which were of high interest at the time of the Boskin Report's publication, and which are associated with technological innovation in ICT equipment and store-to-store substitution for larger stores.

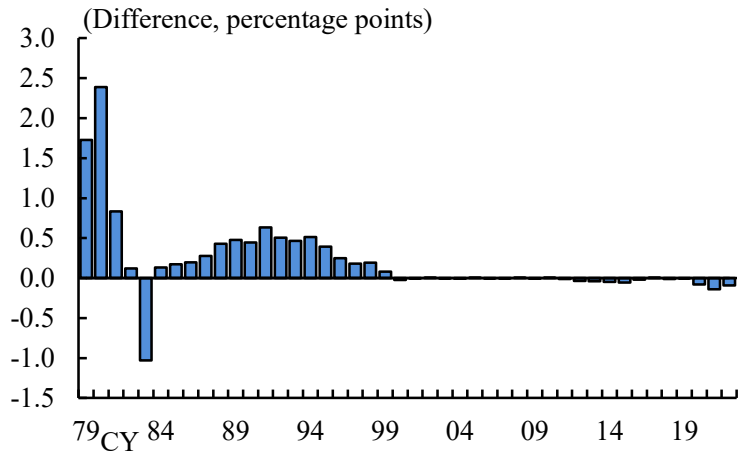
19 Although there are not many quantitative evaluations of new outlet bias in Europe, Covas and Silva (1999), for example, reported a reduction of new outlet bias for the Portuguese CPI.

experts of the BLS (e.g., Abraham, Greenlees, and Moulton, 1998; Deaton, 1998; Diewert, 1998), further validation of the report was conducted to clarify the issues (e.g., Lebow and Rudd, 2003, etc.), and there was the promotion of solutions by the BLS to improve the accuracy of the index. Chart 5 shows a discrepancy between the Retroactive CPI-Urban Research Series using current methods (R-CPI-U-RS), which applies these improvement solutions retroactively prior to the time at which each improvement solution was taken, and the main index. This discrepancy can be interpreted as indicative of the magnitude of the measurement errors that existed at the time of the main index's publication. A review of the time series reveals that the divergence of approximately 0.5 percentage points observed from the late 1980s to the 1990s had disappeared by the 2000s, which suggests that approximately half of the 1.1 percentage point upward bias identified in the Boskin Report may have already dissipated.

Another factor is the change in the economic environment itself. As noted by Gordon (1990) and Greenwood, Hercowitz, and Krusell (1997), the substantial decline in the price of ICT devices was a topic of considerable interest in the 1990s and 2000s, both in the field of statistics and in macroeconomic research as evidence of rapid investment-specific technology progress. However, this downward trend appears to have slowed in recent years. Chart 6 illustrates the price index trends for the representative items of ICT equipment: desktop personal computers, laptop personal computers, and cameras. Although the price index declined rapidly from the early 2000s to the mid-2010s, the downward trend has since slowed, and the growth rate has remained near zero since the beginning of the 2020s. In this context, Takahashi and Takayama (2023) point out that the rate of investment-specific technological progress may have slowed in many countries in recent years. Changes in the pace of technological innovation are likely to have important implications for measurement errors, particularly with respect to the upper-level substitution bias, through their impact on relative price changes.

The following section provides an overview, by type of bias, of the changes in the CPI calculation methodology and the changes in the economic environment that may have affected the magnitude of the bias recently.

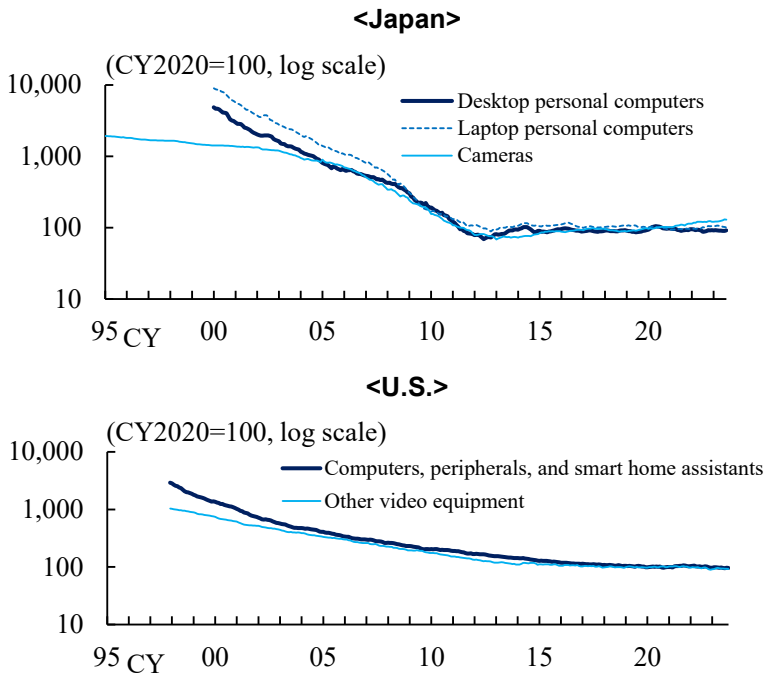
Chart 5 Difference between Main and the Retrospective Index (Measurement Errors Trend)



Note: This chart shows the year-to-year deviation between the main index and the retrospective index. The main index is CPI-U and the retrospective index is R-CPI-U-RS. Both are total CPI (all items). The R-CPI-U-RS is an index in which sequential index accuracy improvement solutions have been retroactively applied to the past CPI-U of the U.S. For example, although PC quality adjustments in the main index have been more advanced since 1998 (use of regression models), the index to which the BLS applied this solution retroactively before 1997, prior to the revision of this solution, is the retroactive index. The difference between the two corresponds to the measurement errors that have been decreased until now. Even if the value of this difference were zero, the CPI-U measurement errors themselves would not be zero when the bias remains in the current calculation method of the CPI-U.

Source: Bureau of Labor Statistics.

Chart 6 End of the Price Decline of ICT Equipment



Note: Japan includes the impact of the consumption tax change. Other video equipment in the U.S. includes video cameras, recorders, etc.

Sources: Ministry of Internal Affairs and Communications; Bureau of Labor Statistics.

(1) Upper-level Substitution Bias (Substitution between Items)

Apart from biases related to the index formulas, the source of the upper-level substitution bias is the inability of statistical authorities to respond immediately to changes in the spending basket due to sudden and substantial demand shifts (including the unavailability of accurate and rapid spending-side statistics). In other words, if the CPI can be updated more frequently, while ensuring accuracy, for the item-by-item breakdown of the basket covered by the CPI, this bias could be reduced.

In this respect, after the Boskin Report, statistical authorities in each country increased the frequency of weighting updates, and these solutions are thought to have contributed to the reduction of this bias. In the U.S., when the Boskin Report was published, the frequency of updates was once every ten years, but after January 2002, it was updated every two years, and then annually from 2023.[20] The European HICP standard made it a rule to update weights annually in 2012. In Japan, weighting changes are typically conducted at five-year intervals. However, since the 2000-base, the "Midpoint-year Revision" has been implemented to assess potential item revisions or eliminations in the midpoint-year between CPI-base years. Chart 7 summarizes the past actions taken in the Midpoint-year Revision in Japan.[21] For example, under the 2005-base, new products such as "beer-flavored alcoholic beverages" were incorporated, while "TV sets (CRT)" was eliminated.

20 At that time, the reason for the infrequent updating of expenditure weights in the U.S. was concern about the accuracy of the values and the time lag in reflecting them in the CPI statistics (CPI Commission, 1997). Based on discussions following the release of the Boskin Report, these concerns were also addressed.

21 Standards for the addition (or elimination) of an item are as follows: Items are added (or eliminated) when the consumption structure changes due to the widespread use of new goods and services or changes in preferences, and when they become more (or less) important in terms of household expenditures. The level of importance is judged on the basis of whether the ratio of the item to household consumption expenditure is, in principle, more than 1/10,000 in the special survey in the "family income and expenditure survey." In addition, items will also be added (or eliminated) in cases where they improve (or can maintain, after deletion of the item) the accuracy and the representativeness of the mid-category index, or where smooth price surveys are possible (or difficult) and price changes can be accurately captured (or is no longer possible).

Chart 7 Details of Review under the Midpoint-year Revision in Japan

Base year	Category and contents	How to reflect on statistics
2005-base	【Adding of items】 (a) "Beer-flavored alcoholic beverages" (b) "Washing & Drying machines" (c) "Video games (portable)" 【Unification of items】 (d) "TV Sets (CRT)" (e) "Media for audio recording" 【Adding of brands/products】 (f) "IP phone charges"	✓Change in item weights[1] (a～c) ✓Weight integration[2] (d～e) ✓Incorporated into an existing item[3] (f)
2010-base	【Adding of brands/products】 Handset cost and telephone charge of smartphone; Handset cost of tablet computer	Incorporated into an existing item[4]
2015-base	【Adding of brands/products】 Heated cigarette; Low-cost smartphone communications charge; SIM-free handset	Incorporated into an existing item[5]

Note: 1. (a): Maintained the weighting of the upper category "alcoholic beverages" and changed the percentage of item weight belonging to this category. (b): The weight of the existing item "washing machines" was split into "(fully automatic) washing machines" and "washing and drying machines." (c): The weight of the existing item "video games" was divided into "video games (stationary)" and "video games (portable)."

2. (d): Existing item "TV set (CRT)" and "TV (LCD)" were unified into "TV (LCD)" (weight integration). (e): Item "media for audio recording" and "recordable DVD media" were unified into "recordable DVD media" (weight integration).

3. (f): Incorporated into the existing item "telephone charges."

4. Incorporated into item "cellular phones," "mobile telephone charges," and "personal computers (notes)" without change in item weighting ("personal computers (notes)" refers to laptop computers).

5. Incorporated into item "tobacco (domestics)," "tobacco (imported)," "telephone charges (mobile phone)," and "mobile phones" without change in item weighting. Some item names were changed from the 2010-base.

6. Based on information published by the Ministry of Internal Affairs and Communications.

Source: Ministry of Internal Affairs and Communications.

In addition to the methodology employed in the calculation of the statistics, changes in economic structure can also influence the magnitude of the bias. For example, the rapid decline in ICT device prices through the mid-2010s, and the subsequent slowdown in the rate of decline, as shown in Chart 6, may have had the effect of first increasing and then reducing the variability of the rate of price change among items. Previous studies mentioned in Appendix 6 point to a positive relationship between the variation in price change rates across items and categories and the magnitude of the upper-level substitution bias, suggesting that the slowdown in the price decline of ICT equipment could have pushed down the upper-level substitution bias.

(2) Lower-level Substitution Bias (Substitution within Items)

The magnitude of the lower-level substitution bias depends largely on differences in price survey methods, and thus there are suggestion that it is not a major problem in Japan and Europe.[22] On the other hand, the number of survey items in the U.S. is smaller than in Japan and Europe, and the U.S. survey method, which uses a probability-proportional-to-size sampling method, is not highly homogeneous within items. As a result, there may be wide variations in the price levels and trends of the price-surveyed brands or products within each item. In general, when arithmetic averages are used when there are large differences among specifications of brands/products, the overall movement of an item index is more likely to be influenced by the price movement of particular brand/product prices. While there is no inherent problem if the price-surveyed brands/products are representative of the overall movement of the item, there is no certainty that this assumption will always be satisfied.[23] In the U.S., prior to 1998, the arithmetic mean (Carli index) was used to calculate item indexes, which the Boskin Report pointed to as a source of the lower-level substitution bias. After this suggestion, the geometric mean has been employed for about two-thirds since January 1999, and GAO (2000) measures this change in calculation method from the arithmetic mean to the geometric mean in the lower-level aggregation as a 0.2 percentage point reduction in the bias of the overall index.[24]

---

22 Yoshizoe (2017) pointed out the same issue. For similar discussions, see the summary of the 14th BOJ-IMES Conference organized by the Institute for Monetary and Economic Studies of the Bank of Japan (Takahashi, Watanabe, and Fujiki, 2007) and Broda and Weinstein (2007).

23 If the share of firm, brand, and product within an item can be identified timely and either weighted appropriately or the number of surveyed brands/products assigned according to their shares, demand shifts can be reflected in the item price indices. However, there are not many cases where market share data is available timely, and a practical solution could be to select an index calculation method that is not excessively influenced by the movements of specific brands/products, while surveying as much as possible brands/products that are considered to be highly representative. As a result, in the context of the lower-level substitution bias, the averaging method used to calculate price indexes is often the focus of discussion.

24 As mentioned above, calculating an item index by geometric mean from brand/product price indices is equivalent to assuming that the price elasticity among brands/products is equal to 1. Therefore, in switching to the geometric mean, items for which there was zero substitution between brands/products (e.g., rent and owner-occupied housing costs, utilities, and medical services) were excluded from application (Lebow and Rudd, 2003). Lebow and Rudd (2003), while noting that there is no hard evidence, pointed out that the price elasticities between brands/products were slightly greater than 1 for the items where geometric averages were applied. They also noted that the elasticities of the geometric mean indexes for items with small samples were slightly below 1, and they pointed out that since the geometric mean index in such cases is biased upward (McClelland and Reinsdorf, 1999), the lower-level substitution bias (mainly for items with a small number of brands) remains after the geometric mean transition.

In Europe, the quality and price differences between specifications (brands/products) within an item index are not expected to be large due to the large number of items and the price surveys that specify representative stores and the survey characteristics of the brands and products. Furthermore, in the HICP, the harmonized European standard, 11 countries accounting for 63% of the expenditure share use the geometric mean, and 7 countries use the arithmetic mean (ECB, 2021), which means that more than half the countries use a formula that is less likely to have the lower-level substitution bias.[25]

Japan also has a large number of items, which is expected to ensure the homogeneity of surveyed specifications within an item. In Japan, a survey method is adopted that makes it easier to fix the quality within an item, such as specifying the brand/product name by item or, if not specified, defining the capacity, packing method, etc., in detail. Therefore, it is said that the level of lower-level substitution bias is low and almost does not need to be considered (Shiratsuka, 1999). In order for the surveyed brands/products to be representative of the entire item category, they must be replaced by new products in a timely manner, and in this regard, they have been replaced almost monthly since 2004. However, even in such a case, as Shiratsuka (2023) points out, there could be another problem, namely whether the surveyed brand might be representative of the movement of the item.[26]

25 The ECB (2021) measures the upper- and lower-level substitution bias for food products. In European price indexes that use similar survey methods, Germany, which uses a ratio of average prices at two points in time (arithmetic mean, Dutot index), has a larger lower-level substitution bias than France, which uses a geometric mean of price ratios between two points in time (Jevons index), suggesting that differences in index formulas have a larger impact on the item index.

26 For example, only "chocolate bars" prices are surveyed for the item "chocolate" where the brand name is specified (one brand per item survey). Although the replacement of surveyed brands from time to time maintains the representativeness of the surveyed chocolate bars (trademarks, capacities, etc.), it is not clear whether the price trends of chocolate bars can represent the movement of the entire "chocolate" product category. Shiratsuka (2023) compares CPINow, calculated using supermarket POS data, and CPI to examine how brand-specific price movements affect the item index. The study points out the possibility that the surveyed brands may not be representative of the overall movement of the item index, as the movements specific to the surveyed brands diverge from those of other samples in the case of chocolate, etc. (items with a high degree of brand specification) and sausages, etc. (items with a low degree of brand specification).

(3) Quality Change and New Product Bias

(Quality Change Bias)

As noted above, many studies of measurement errors from the late 1990s through the 2000s pointed to the contribution of quality change bias as the primary cause of the overall upward bias. At the item level, prices of consumer durables such as personal computers and camcorders had been declining. However, as pointed out by Shiratsuka (1995a) and others, the lack of quality adjustment in these price declines was thought to be the cause of the upward bias in the overall CPI.[27] For example, Bils and Klenow (2001) point out that while the quality of U.S. consumer durables improved at an annual rate of about 3.7% during 1980-1996, only 40% of the improvement was quality-adjusted, and the remaining 2.2% portion remained as an upward bias in the price index for consumer durables.

In response to these suggestions, statistical authorities in various countries have systematized and refined their quality adjustment methods, such as by expanding the scope of application of hedonic methods, and as a result, the quality change bias seems to have decreased over the past 20 years. For example, in GAO (2000), which was a reassessment of measurement errors by the Boskin Report authors, all of the authors found that the use of hedonic methods introduced by the BLS for PCs and TVs affected the magnitude of the quality change bias.[28]

Chart 8 summarizes the items to which the hedonic method applies in the CPI of the U.S. and Japan. For ICT devices (PCs and digital cameras), which have been pointed out in previous studies as having a quality change bias, the hedonic method was introduced sequentially (in Japan, after the year 2000; in the U.S., PCs in 1998, TVs in 1999, and camcorders in 2000). In addition to ICT devices, the hedonic method has been applied to a wide range of goods and services in the U.S., including white goods, clothing, and internet services.

27 Quality change bias for consumer durables in Japan was discussed in detail by Shiratsuka (1995a) for PCs, Ota (1980) and Shiratsuka (1995b), etc. for automobiles, and Shiratsuka and Kuroda (1995) for video cameras.

28 GAO (2000) reported an overall change in measurement errors of -0.3 percentage points due to the change in statistical procedures.

Chart 8 Examples of Items to Which the Hedonic Method Is Applied in the Japan and U.S. CPI

	Major items	Japan	U.S.
Goods	Televisions	○	○
	Desktop personal computers	○	△Note
	Laptop personal computers	○	△Note
	Cameras	○	○
	Phones, accessories, and smartwatches		○
	Video equipment		○
	White goods		○
	Clothing; Footwear; Watches		○
Services	Rent and owner-occupied housing costs		○
	Telephone services; Internet access services		○
	Cable and satellite television service		○

Note: The hedonic method was used for the PC of the U.S. CPI, but is now changed to the cost-based adjustment method. Based on information published by the Ministry of Internal Affairs and Communications and Bureau of Labor Statistics.

Sources: Ministry of Internal Affairs and Communications; Bureau of Labor Statistics.

Nevertheless, the growth rate of the price index does not always shrink when the quality adjustment method is switched to the hedonic method. Brown and Stockburger (2006) noted that the measurement errors regarding quality adjustments and new products can bias both up and down, depending on the item-type and the time. As an example, they stated that the U.S. hedonic method for clothing had either an upward or a downward impact, depending on the point in time and the category. Johnson, Reed, and Stewart (2006) compared price indexes calculated using hedonic and conventional quality adjustment methods (Chart 9). By segment, computers and peripheral equipment has a consistent quality improvement trend; however, the existing quality adjustment method could not fully capture the quality improvement, resulting in an upward bias. In contrast, for rents, housing quality deterioration was valued in the direction of pushing the index up by hedonic methods. Lebow and Rudd (2003) calculated the breakdown and contribution of changes in quality change and new product bias in 1996, when the Boskin Report was published, and in 2001, after several solutions were implemented to improve

accuracy (Chart 10).[29] The "(a)-(b)" in the right column corresponds to the difference in contribution to the overall index, with a negative value indicating a reduction in quality change and new product bias. The results of this analysis suggest that about half of the upward quality change and new product bias that existed in 1996 had disappeared by 2001 (quality change and new product bias at the time of the Boskin Report was +0.69 percentage points (recalculated value, based on 2001 weights), which had declined by 0.32 percentage points by 2001). By category, the upward bias was decreasing in education and communications, which includes personal computers and peripherals, recreation, which includes televisions and audio equipment/other video equipment, and apparel, while the upward bias in the medical sector remained high at +0.17 percentage points.

Chart 9 Impact of Adoption of Hedonic Method on the CPI Bias (U.S.)

	After hedonic adjusted index − old index (annual rate, percentage)
Information processing equipment	**-3.81**
Computers	-6.50
College textbooks	-2.53
Washers	-0.78
Microwaves	-0.17
Televisions	-0.11
Refrigerators	+0.02
Dryers	+0.06
Camcorders	+0.15
Housing	**+0.31**
Apparel	**+0.39**
Audio equipment	+1.52
VCR's	+1.89

Note: This chart indicates the percentage of deviation in each category, not the contribution to the total CPI (all items). If the value is negative, it means that applying the hedonic method to the item index resulted in either a slower increase in the hedonic adjusted index or a more rapid decrease in hedonic adjusted index compared with the previous quality adjusted index.

Source: Johnson, Reed, and Stewart (2006).

29 The change in the quality change bias from the Boskin Report in Lebow and Rudd (2003) includes the effect of revisions based on the results of new studies published after the publication of the Boskin Report, as well as changes in the BLS quality adjustment methodology.

Chart 10 Decrease in Bias since the Boskin Report

	(a)2001		(b)1996	(a)-(b)
	Bias	Cont. to total	Cont. to total	Diff. of cont.
Medical care	**+2.3**	**+0.17**	**+0.21**	**-0.04**
Medical care services	+2.5			
Prescription drugs	+1.2			
Education and communication	**+1.0**	**+0.06**	**+0.10**	**-0.04**
PC services (Internet)	+19.0			
PC and peripherals	+4.0			
Housing (tenants' rent and owners' equivalent rent, etc.)	**+0.1**	**+0.04**	**+0.07**	**-0.03**
Food	**+0.2**	**+0.03**	**+0.05**	**-0.02**
Recreation (televisions and video equipment, etc.)	**+0.3**	**+0.03**	**+0.09**	**-0.06**
Transportation (vehicles and motor fuel, etc.)	**+0.1**	**+0.01**	**+0.04**	**-0.03**
Apparel	**+0.0**	**+0.00**	**+0.07**	**-0.07**
総合		**+0.37**	**+0.69**	**-0.32**

Note: Units are percentage points. Estimates by Lebow and Rudd (2003). Calculated quality change and new product bias for each category. The contribution ratio to total index in (b) was recalculated using the weights at the time of comparison and the Boskin Report's original estimates. Copyright American Economic Association; reproduced with permission of the Journal of Economic Literature.

Source: Lebow and Rudd (2003).

(New Product Bias)

It is often pointed out that the new product bias will widen, especially when it is not possible to reflect in the CPI price declines that occur early in the product cycle immediately after the launch of a new product, and when it is not possible to assess quality differences between new and old products.[30] In this regard, the above-mentioned change in the frequency of review of CPI items and the application of the hedonic method to a wide range of items may have contributed to a reduction of the new product bias.[31] In particular, Lebow and Rudd (2003) pointed to the potential use of external data, such as

30 The Boskin Report also provides an example of a room air conditioner first sold in 1951 that was not included in the CPI until 1964, taking 13 years to be included.

31 Shiratsuka (1998) pointed out that the method to accurately capture quality changes (such as the hedonic method) was not adopted, which may have led to a delay in item inclusion for products with short product cycles, such as ICT equipment. In the absence of an appropriate quality adjustment method, if goods with short product cycles are selected as survey products, it can result in an item index with almost no fluctuation, as it may not reflect quality adjusted price changes when switching between old and new products, which occurs frequently.

scanner data, which have more brands and products than those covered by the CPI, and that such data could also be useful for tracking the release of new products.

(4) New Outlet Bias

For Japan's CPI, changes in the method of calculating the statistics and changes in the structure of the economy are thought to be working in the direction of reducing new outlet bias. First, as pointed out by Shiratsuka (2006, 2007), the method of selecting stores for the survey was changed[32] in 2003. Specifically, the method is based on dividing the survey areas according to the existence of regional differences in the prices of goods and physical and geographical boundaries, and then surveying representative stores within these areas.[33] Goods sold in representative stores are usually considered to be the cheapest, and new outlet bias is expected to decrease as survey outlets are changed from time to time. This survey method is consistent with the idea that when people travel across rivers, mountains, etc., transportation costs and opportunity costs associated with moving around increase, and that business areas are divided between adjacent survey areas.[34]

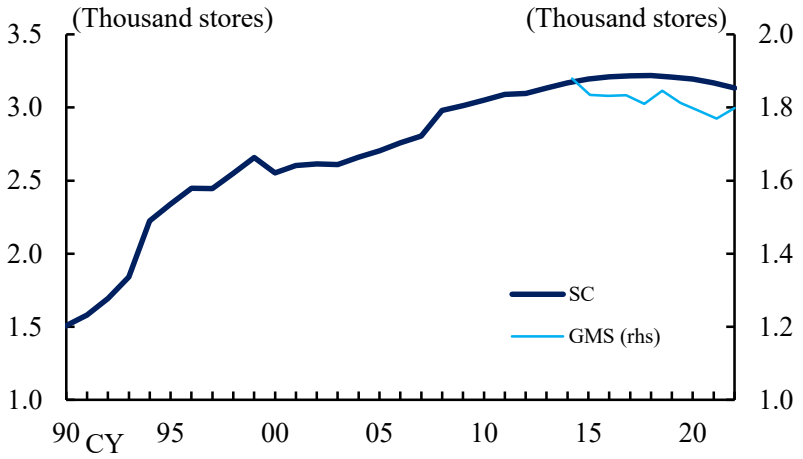
32 Technically speaking, the method of setting survey districts in the Retail Price Survey, which is the basic statistic of the CPI, was changed and this new rule has been implemented from the January 2004 survey onward. In terms of the base year of the CPI, this methodological change in the definition of the survey areas corresponds to the minor revision in the middle of the base year 2000 (major benchmark year revision took place in August 2001).

33 According to the Ministry of Internal Affairs and Communications (2003), the following three steps are taken: (1) classify items into broad categories according to whether they are more frequently purchased in residential or urban areas and whether there are price differences between districts; (2) divide price survey districts based on the geographical distribution of commercial concentration areas and topographical features (rivers, railroads, etc.); and (3) assign the number of survey prices (number of brands/products) to each survey district of (2) according to the characteristics of the items identified in (1) above. Price surveys are then conducted in representative stores within each district.

34 The method of setting detailed locations of surveyed stores for each goods item is also consistent with the implications derived from the "Hotelling model," a model of industrial organization theory based on location competition (horizontal differentiation). Assume a Hotelling model with two stores providing the same good in a region where consumers are evenly distributed along a linear line. If the parameter related to the cost of travel for a consumer residing between the two stores is set to a sufficiently large value, the trade area is separated between the two stores, and each store has price dominance in each of the separated trade areas. Therefore, from the perspective of the model representing business competition, it is a natural assumption that the market is considered disconnected, and the survey area is divided by geographic boundaries where monetary, time, and psychological travel costs increase significantly. Incidentally, as the parameter on travel costs is reduced closer to zero, goods sold in two stores are considered more substitutable, and equilibrium prices in such markets (trade areas) converge to marginal costs, which is equivalent to the equilibrium price given by the Bertrand model.

In addition, the economic structure of the retail sector could had changed in ways that make new outlet bias less likely. As shown in Chart 11, from the mid-1990s to the early 2000s in Japan, general merchandise stores (GMSs) and shopping centers (SCs) aggressively expanded their store openings, while offering goods at relatively low prices, and consumers shifted their shopping channels from individual mall stores to GMSs and SCs. Although these changes created an environment conducive to new outlet bias[35], the number of GMS and SC stores had leveled off since the 2010s and has declined slightly in recent years.

Chart 11 Number of GMS and SC Stores in Japan



Sources: National Supermarket Association of Japan; Japan Council of Shopping Centers.

It has been noted that, in the U.S., substitution across store types could have affected new outlet bias in the retail market. Greenlees and McClelland (2011) find that from 2002 to 2007, discount department stores and warehouse club stores increased their share of CPI survey prices, that these stores were significantly less expensive than large grocery stores, and that the resulting substitution between stores contributed to the decline in

35 Higo and Shiratsuka (2023) and Shiratsuka (2023) examine time-series changes in new outlet bias based on POS data on food and beverage products sold at retail stores in Tokyo. The studies report that the new outlet bias persisted from the mid-1990s to the early 2000s, noting the significant special sales by large stores and the associated significant interstore substitution effect. After that period, the bias has remained somewhat smaller, with ups and downs.

average prices surveyed by the BLS.[36, 37]

Incidentally, as a statistical solution to new outlet bias, various quality adjustment methods have been proposed since immediately after the Boskin Report; however, at present, quality adjustment for price differences when survey outlets are changed has only been partially implemented in some countries.

36 A similar analysis by Hausman and Leibtag (2009), using household data, found that food prices at new types of outlets, such as supercenters and club centers, which expanded rapidly in the 1980s and 1990s, were 27% lower than those at traditional outlets. Nevertheless, Greenlees and McClelland (2011) identify much of this price difference as a difference in the quality of service between stores.

37 Lebow and Rudd (2003) reduced the new outlet bias from the Boskin Report value to +0.05 percentage points, which was the result of choosing the median rather than the upper limit of the previous study that was originally referenced. Adding the contribution from online purchases to Hausman and Leibtag's (2009) estimate of +0.32 percentage point upward bias in CPI food price growth due to the outlet bias, Moulton (2018) estimates an upward bias of +0.08 percentage points.

5. Remaining and New Issues

Given the accumulation of research conducted since the publication of the Boskin Report, efforts to improve the index in each country, and changes in the economic structure, it appears that measurement errors have generally decreased with respect to the issues raised in the Boskin Report, although there are differences between countries and among types of bias. Nevertheless, it does not mean that the CPI and COLI are equal at this time. In this section, as "remaining issues," we first highlight the discussion of measurement errors in several service prices. Next, as "new issues," we introduce the discussion of the implications of new commercial channels, such as electronic commerce (e-commerce), which did not exist at the time of the Boskin Report's publication.

(1) Price Survey and Quality Adjustment of Service Prices

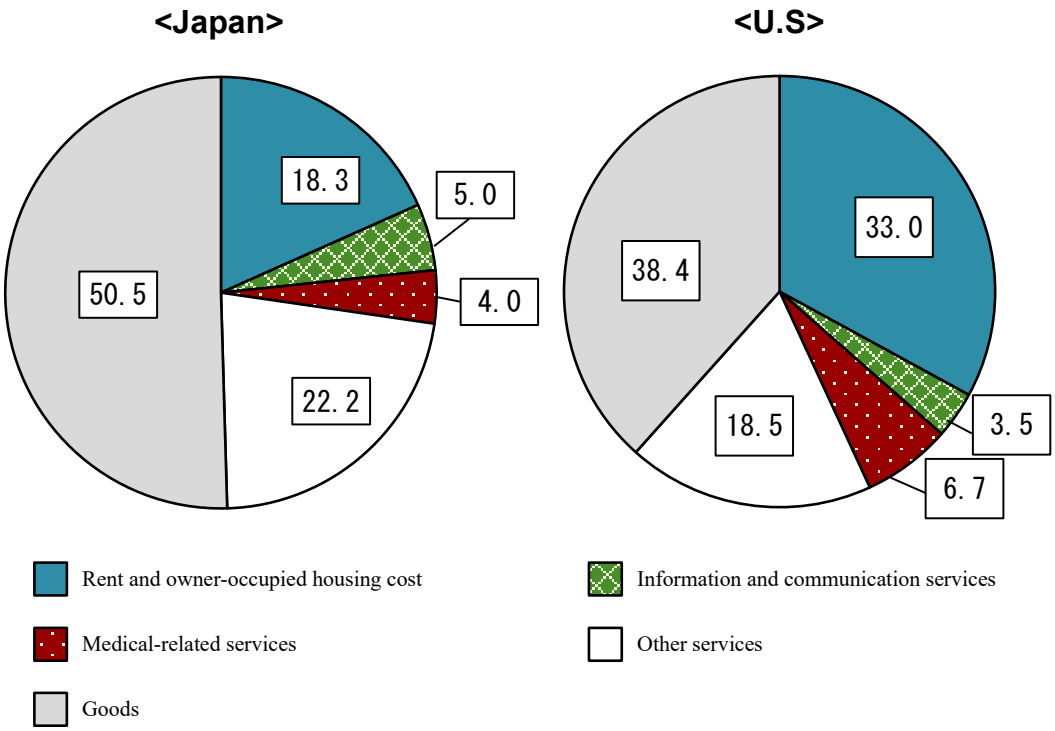
Regarding issues in price surveys for services, there are difficulties derived from the characteristics of services, such as survey coverage and quality adjustment problems, as well as the fact that the trend of digitalization is shifting toward services. The former issue, more specifically, is that the service is provided with various contents and conditions reflecting the various preferences of customers,[38] and that the quality varies depending on factors that cannot be visualized, such as the skill of the service provider. These characteristics are considered to make it difficult to identify the type and content of services to be surveyed and to apply quality adjustment methods such as hedonic methods. For the latter issue, changes in the economic structure, such as the shift in digitalization from goods to services, are likely to cause difficulties in quality adjustment and price surveys using existing methods.

In this section, we first present recent discussions on "rent services, etc.," "information

38 One of the ways to characterize services is the concept of "vertical differentiation" and "horizontal differentiation." In vertically differentiated markets, firms seek to differentiate themselves vertically from other firms' products in terms of quality (specifications) and price range. Since each quality can be ordered vertically, for example, the prices of appliances with exactly the same specifications are expected to be roughly at the same level. On the other hand, in the market where horizontal differentiation takes place, each consumer has different preferences and demands, therefore, firms differentiate their products according to their preferences (products are differentiated horizontally). In both types of differentiation, if the characteristics of the goods or services can be identified, quality adjustment should in principle be possible, e.g., Lancaster (1966), Rosen (1974). However, in the case of horizontally differentiated markets, consumers who derive utility from certain characteristics and consumers who do not derive utility at all can coexist. Therefore, when there is heterogeneity in consumer preferences in a market with horizontal differentiation, it is not easy to identify the relationship between price and quality with the hedonic function used in the quality adjustment of price statistics, even if all characteristics are known.

and communication services," and "health care-related services," based on their weights and magnitude of contribution among the services. Next, we mention an attempt to measure service prices in the digital sector.

Chart 12 Basket of the CPI



Note: Japan weights are as of 2020-base and the U.S. weights are as of 2021-base. Medical-related services include medical fees, nursing care fees, and accident insurance fees, etc. Information and communication services include communication charges (fixed-line and mobile phones), Internet access charges, and broadcast reception charges, etc.

Sources: Ministry of Internal Affairs and Communications; Bureau of Labor Statistics.

(Owner Occupied Housing Cost)[39]

Chart 12 shows the CPI basket, which is expenditure weight. Rent, including owner-occupied housing (OOH) cost, accounts for 20% of household expenditures in Japan and

39 The quality change bias in rent and owner-occupied housing cost was discussed in the Boskin Report, which indicated that there was an upward bias of about +0.25 percentage points between 1976 and 1996 for the index of rent and owner-occupied housing cost (shelter). The report suggested that the reason for this upward bias was not the aging of the house, but rather the fact that it does not reflect the improved quality of the appliances installed in the rooms. This suggestion appeared to be due to the fact that (although the report was skeptical about the fact that quality adjustment by BLS captured age-related deterioration), quality adjustment using the hedonic method had already been performed since 1988 for age-related deterioration, etc., but the quality adjustment of the facilities, such as air conditioners and bath piping, had not been conducted.

30% in the U.S. (16% and 25%, respectively, for the OOH portion), and the utility derived from the OOH is likely to be a large part of the total utility that households receive from various consumption. Therefore, although many believe that the OOH costs should be included in the coverage of general price statistics, there are currently cases where the OOH costs are excluded from the official series of the price index, such as the HICP, which is the ECB's targeted price index. There is no international consensus on the measurement method of the OOH cost, partly because it is not actually traded, and prices cannot be directly observed.

Based on IMF *et al*. (2020), the calculating methods of the OOH cost can be roughly classified into (1) "use approach," (2) "payments approach," and (3) "acquisitions approach" (Chart 13). The three approaches are based on different measurement ideas and capture, respectively, changes in the cost to households of consuming the same services as OOH, changes in repayment amount related to OOH, and changes in the costs associated with acquiring and maintaining housing. The "rental equivalence approach (owners' equivalent rent of residences, OER; imputed price)," which is used in Japan, the U.S., the U.K., etc., is classified as the use approach and is a measurement method that estimates the cost of OOH-related services that households consume at the market price of similar housing (rental house rent). A relatively large number of countries have adopted this approach, although there are requirements for adoption, such as the need to use rental data and the existence of a rental market with sufficient depth and transparency.[40] The payments approach used by Ireland and other countries uses mortgage repayments and insurance amounts, and the acquisitions approach adopted by Australia and New Zealand uses acquisition costs and related costs for new properties. For the former payments approach, there is the problem that the amount of loan repayments is often affected by the monetary policy interest rate at the time. For the latter acquisitions approach, there are

40 In order to use the rental equivalence method, rents for rental properties in the same or comparable locations as owner-occupied houses need to be surveyed, but this assumption may not be satisfied if the size of the rental market is small (ECB, 2021). In addition, in some countries, there are long-term rent contracts and regulations that protect sitting tenants from fluctuations, which can create a problem of disparity between the rents paid by existing tenants and those paid by new tenants.

issues such as whether it is easy to obtain data separating housing and land prices,[41] and how to separate the cost of the utility portion of housing services from the cost of the housing investment (which corresponds to future utility).[42]

Chart 13 Classification of OOH Cost Calculation Methods

Category	Examples of survey method	Adopting country
Use Approach	Rental equivalence[1]	Japan the U.S. the U.K.
	User cost[2]	Canada Sweden
Payments Approach	Loan repayments and insurance payment, etc.	Ireland
Acquisitions Approach[3]	Acquisition cost of new properties and related cost, etc.	Australia New Zealand

Note: This classification is based on IMF *et al*. (2020).

1. Owners' equivalent rent. Households produce and consume on their own the services they benefit from owner-occupied housing, and this method evaluates the price of these services at market prices (rents for rental housing).
2. A method of evaluating OOH cost based on loan repayments, insurance, depreciation, opportunity cost of living, and capital gains, etc.
3. Land is considered an investment element, while buildings are considered a consumption element since they depreciate with consumption, and the latter and its associated costs are included in the aggregate (IMF guideline).

In Europe, OOH costs have not been included in the HICP due to the lack of consensus on the treatment method of these costs across EU countries and the difficulty of obtaining

41 In the acquisitions approach, land is considered as the investment element and buildings as the consumption element, which deteriorates with consumption, and the latter and its associated costs are considered for CPI aggregation (IMF guidelines). The Owner Occupied Housing Price Index (OOHPI) published in Europe adds taxes and fees related to real estate acquisition to the cost of new construction (including self-build and major renovations) and calculates them at the time of purchase. It has been pointed out that the use approach is more consistent with economic theory than the acquisition approach in terms of measuring the value of owner-occupied services at that point in time (Diewert and Nakamura, 2009). On the other hand, it has also been pointed out that other durable goods other than housing are generally measured using the acquisitions approach, which would ensure consistency with other items if OOH cost is measured using the acquisitions approach.

42 Given the large share of owner-occupied houses in many countries, which approach is adopted could also affect the overall CPI. For example, Cecchetti (2007) points out that if the BLS had used house prices instead of rents to calculate the price index for owner-occupied consumption in the U.S. after 2000, the core CPI would have been about +1.5% higher than the actual CPI on an annual average basis. The Australian Bureau of Statistics (2023), which uses the acquisitions approach, found a high correlation between OOH costs and the Producer Price Index (PPI, house construction costs including labor costs) in Australia, and notes that, since the spread of the pandemic, the OOH cost had increased significantly relative to other countries using other calculation methods, as a result of high material prices (lumber, steel, concrete, etc.) and construction demand.

source data. However, the "Strategic Review" by the ECB (ECB, 2021) announced the ECB's plan to include OOH costs in the HICP and that it had been decided OOH costs would also be taken into account in monetary policy decisions. This decision was due to the significant impact of housing costs on people's perception of inflation, the comments from the general public that housing costs should be reflected (ECB, 2021; Eiglsperger *et al*., 2022), and the high percentage of people who have OOH in Europe[43] (Eiglsperger *et al*., 2022). In the U.K., while the CPI referenced by the Bank of England (BoE) does not include OOH cost, the "CPIH,"[44] which the Office for National Statistics (ONS) has positioned as a main index since 2017, does includes OOH cost, which means that the central bank and the statistical authority place different emphasis on different indicators.[45]

When using the rental equivalence approach, quality adjustment is not straightforward, due to the large number of rent quality determinants and their complex combinations. In Japan, studies on rent quality adjustment have been accumulating; for example, in response to issues raised by the Statistics Commission in the Ministry of Internal Affairs and Communications, Onaka (2022) and the Ministry of Internal Affairs and Communications (2021a) conducted analyses using data from the Housing and Land Survey's questionnaire for households living in rented housing. As a result of these studies, the rate of quality change in rental housing is estimated to be -0.8 to -0.9% per year for wooden houses and -0.7% for non-wooden houses, both of which are identified as being in the direction of deterioration over time. Applying these values to the sample distribution of the CPI, the year-on-year change in the quality-adjusted private rents index corresponds to an upward shift of +0.8 percentage points for wooden houses and +0.7 percentage points for non-wooden houses, compared with the pre-quality-adjusted index (approximately +0.1 percentage points in terms of contribution to the overall index).

43 Regarding measurement methods, it appears that either the acquisitions approach or the use approach (rental equivalence method) is being considered, because if the payments approach is adopted, the base data for the calculation will include loan repayments, which are affected by policy interest rates (ECB, 2021; Eiglsperger *et al*., 2022). In comparing the two approaches, it is argued that the acquisitions approach raises concerns about excessive OOH weight fluctuations in some countries with pronounced construction cycles, and that the rental equivalence method raises concerns about the depth of the rental market, etc. (ECB, 2021).

44 The ONS website states "CPIH is the most comprehensive measure of inflation."

45 The CPIH in the U.K. was released in 2013 as a government priority; however, it was removed from the "National Statistics" status in 2014 after a calculation issue was discovered (Bean, 2016). The different views on the robustness of the statistics might be the reason for the difference in the positions of the BoE and the ONS.

One study points to a downward bias resulting from deviations from marginal rents (Yoshida, 2023). In Japan, the CPI is currently estimated based on the "average" of private rents. However, the OER of OOH is conceptually the opportunity cost of OOH, i.e., "the rent that would be paid if the OOH you are currently living in were newly leased in the rental market." Therefore, it would be more appropriate to relate the OER to the "marginal rent," which is the rent charged for a newly rented property, rather than to the continuing rent. Yoshida (2023) estimates a quality-adjusted marginal rent index that controls for age-related depreciation and other factors by using the propensity score matching and hedonic methods on monthly data (including rent/price, space, location, building age, etc.) for rental and sales properties listed on real estate and housing information websites. This paper finds that the OER index of the CPI is estimated to be approximately 1.3 percentage points lower than the marginal rent index on an annual basis, and that approximately 1 percentage point of this difference is explained by depreciation over time.

In the U.S., quality adjustments to account for change in quality due to renovation, aging, etc. were already implemented at the time of the Boskin Report, but another issue about the survey method has been raised in recent years. The U.S. rent survey is conducted semi-annually for the same tenant, the rents for each month are weighted averages, and the growth rate between the two points in time (six months) is converted to a monthly basis. One advantage of this method is that the same tenant is surveyed, thus ensuring that the characteristics of the rental house remain consistent. However, it has been pointed out that a selection bias can arise due to the exclusion of samples that cannot be continuously surveyed at two points in time (Yoshida, 2023). In other words, because the sample is biased toward continuously cooperative tenants, the composition of ongoing rents tends to increase more than that of new rents.[46]

(Cellular Phone Communication Charges)

Cellular telecommunications services were not adopted as an item in the U.S. CPI at the time of publication of the Boskin Report (Hausman, 1999), and the report positioned them as a service that may be generating a new product bias (upward) rather than as a quality change bias. Since then, in many countries, including Japan and the U.S., cellular

46 Ambrose, Coulson, and Yoshida (2015, 2018, 2023) discuss some measurement errors associated with this survey method, noting, for example, that it tends to under (over) estimate rents during periods of economic expansion (recession) as a result of a bias toward ongoing rents. It is also suggested that there is a downward bias in the CPI due to the inability to capture new rents when tenants change (Lane, Randolph, and Berenson, 1988; Gordon and vanGoethem, 2007; Crone, Nakamura, and Voith, 2010). See Yoshida (2023) for previous studies discussing these points and other issues.

phone communication charges have been adopted as CPI items, etc., therefore, the bias of cellular phone communication charges, in terms of the source of new product bias, can be considered to have been eliminated. However, in light of the diversity of service offerings resulting from providers' differentiation strategies and the variability in customer demand, discussions have shifted from the initial focus of the Boskin Report to different considerations, including the optimal price survey method and quality change bias, such as the development of quality adjustment methods.

In Japan, a model price survey (model formula) is used.[47, 48] Following the classification of service quality and users into multiple categories based on data traffic, call time, and other relevant variables, the price index is calculated by assuming that the most economical plan among the plans in each category is preferred. While this assumption is consistent with the concepts of economic rationality, in reality, the assumption can diverge from reality because it is possible that not all consumers will choose the cheapest plan, or that they will choose it but the transition will be gradual.

Nishimura and Higo (2022) point out the impact of this cheapest plan assumption on the measurement errors. The analysis compares the movement of the price index at the time each telecom company launched its low-cost plans and these plans were reflected in the CPI (April 2021) with the monthly average revenue per unit (ARPU) reported by the major telecom firms and the amount spent on telecommunications in the Family Income and Expenditure Survey. This comparison shows that while it was assumed at the time the statistics were compiled that consumers would shift quickly to low-cost plans, the actual shift was limited, which may have led to a short-term downward bias in the CPI.[49]

47 The "model price survey" here means that the "model" is adopted in the sense that consumers with each preference choose the lowest priced plan among the service plans that match their preferences. Therefore, it is not the virtual price that is the subject of the price survey, but the price at which the transaction actually takes place (in the case of the telecommunications sector, the list price).

48 In addition to the categories of carriers, brands, and telecommunication types, there are multiple categories corresponding to combinations of calling time and data traffic volume, and the cheapest plan in each category is selected as the survey price. When the content of the cheapest plan is changed by the service provider, the "direct comparison method" is often used to connect the old and new prices, assuming that there is no difference in quality within the same category. However, this approach is justified only if the quality-determining factors can be fully identified and the preference categories in the model prices can be accurately defined. There could be cases where this assumption is not valid, such as when quality changes occur between the old and new plans in each category.

49 A possible explanation for the slow or no transition to the lowest priced plan is the switching costs. Specifically, switching costs are considered to include termination fees associated with multi-year contracts, opportunity costs caused by being held for long hours when changing plans, search costs when comparing old and new plans, and psychological costs associated with moving from a familiar provider.

The U.S. CPI covers plans that include voice calls, text messages, and data communications, and, as with other items, uses sampling method to select price-surveyed plan. The Consumer Expenditure Survey (CES) was used to select survey targets (distributors), and the selection of surveyed plans was sampled using a sampling with probability proportional to size based on market share, which to some extent is supposed to reflect the actual market share of each plan. In addition to rotating the surveyed sources every four years, third-party data are also used to complement the dynamism that cannot be captured by the CES and other surveys. When the survey plans are switched to new ones, the BLS analyst determines the quality change, and if the quality is the same, the "direct comparison method" is used; if the quality changes, the "hedonic adjustment method" is used to adjust the quality between new and old telecom plans.

(Health and Medical Care Services)[50]

Medical services also reflect the diversity of disease types and service recipients, making quality fixing, quality adjustment, and survey price selection for survey targets difficult. In addition to the large proportion of the current expenditure weighting, there is the possibility of further increases in health and medical care-related expenditures as the population ages, and there is reportedly a growing demand to refine the deflator used in the SNA calculation.[51] Medical fees in Japan are captured in the model price survey, which is representative of the medical treatment in each of the categories when allocated treatments by age (general, elderly, late elderly) and by the type of treatment (inpatient,

50 Quality change bias in health and medical services was also mentioned in the Boskin Report. The report noted that there are quality changes related to treatment efficacy and other broadly defined quality changes. At the time, price surveys of medical services were tied to the type of service rather than the type of disease, such as the cost of a one-day stay in a hospital room, etc., and the survey did not capture the cost of treatment for the disease, resulting in an upward quality change bias (Lebow and Rudd, 2003). Since 1997, after the publication of the Boskin Report, quality adjustment (quality fixing) in a broader sense has been implemented, such as by beginning to survey the total cost of treatment for specific diseases, and some of the quality change bias appears to have been reduced to some extent. However, even today, quality adjustment for "quality of medical care" in the sense of the effect of treatment on survival rates has not been implemented in statistical practice, and in this context, quality change bias is considered to remain in place.

51 Nishizaki and Kuwahara (2023) note that, while in the U.S. and in academic research, there is a trend toward quality adjustment and disease-specific price surveys on the deflator side, in Europe there is a trend toward adjusting these on the output side.

outpatient (hospital or clinic), dental), and a similar method is employed in the U.S.[52] These categories are small compared to the number of disease classifications. For example, the World Health Organization (WHO)'s International Classification of Diseases (ICD-10) has about 14,000 disease categories, and the next generation classification (ICD-11) will increase the number to about 18,000.[53] If the cost of treatment for each disease is different, the overall price movement may not be adequately represented only by the prices of typical treatments that fall into the above categories. In addition to these diversity issues, there is also an argument that quality improvements resulting from improved medical technology, such as increased survival rates, should be discounted from the price.

Disease-specific price surveys are being considered in the U.S., and pilot indices are starting to be published.[54] Bradley (2013) estimates a price index calculated for each disease based on publicly available databases, and notes that it was 1 percentage point per year lower than the existing price index. The BLS has also calculated and published on a pilot basis disease-based price indexes (DBPI), which are aggregates of price indexes created for each of more than 100 diseases (Chart 14). In addition to the accumulation of price indices for each disease, the volume-adjusted indices capture changes due to technological advances in the use and frequency of medical supplies and medicines during treatment, while other indices are published with adjustments such as the allocation of medication to patients with complications, according to the percentage of treatment for each disease. The index trends show that the volume-adjusted series is lower than the fixed-weighted series, in which the basket of medical treatments is fixed at the base period, suggesting that the number of medicines, etc., needed to treat diseases is decreasing, i.e. that prices, which are total treatment expenditures, are decreasing.

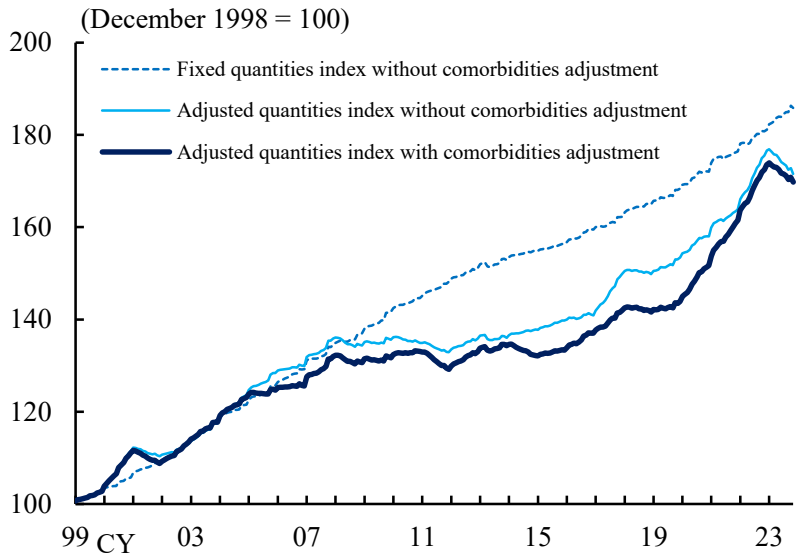
52 Similar to mobile phone communication charges, Japan's model price survey uses a methodology that assumes a set of representative consumers and continuously surveys the prices of major medical procedures seen by these consumers. In the U.S., the most popular health care services are surveyed on an ongoing basis at each health care facility. Classification is based on payment method (self-pay, private insurance, Medicare) and medical treatment category (general medical, dental, ophthalmology, etc.), and to a large extent, the U.S. price survey methodology is similar to that used in Japan.

53 In Japan, this disease classification is used for statistical surveys under the Statistics Act, and for management of medical records in medical institutions.

54 There are suggestions that the CPI in Japan should be subdivided by disease or injury (Nishizaki and Kuwahara, 2023; Nishimura, Yamasawa, and Higo, 2020), but approaches are not as advanced as those in the U.S.

As an example of incorporating quality improvement, Matsumoto (2021), calculating hospital service price indices, performs quality adjustment using information on patient satisfaction and probability of death by condition, finding that between 2010 and 2016, the growth rate of the quality-adjusted price index (PPI) for hospital inpatient services is up to 0.26 percentage points lower than the unadjusted index. Other pioneering work, although not quality-adjustment methods, includes Cutler *et al.* (1998), which assumes a utility function that includes utility from health status and utility from nonmedical expenditures and calculates the price of medical services to have the same utility (COLI). This study focuses on acute myocardial infarction.

Chart 14 the U.S. Price Index Aggregated by Disease



Note: 1. Fixed quantities index is the Lowe index with quantity fixed at the base period. Adjusted quantities index is adjusted for changes in the medical treatment basket (weights are updated on an annual basis).

2. Index with comorbidities adjustment is a prorated index when medical treatments are provided for more than one disease. For example, when a medical procedure requires two treatments for heart disease and one for diabetes, 2/3 and 1/3 are allocated to the heart disease and diabetes indexes, respectively.

3. All indices are not quality adjusted.

Source: Bureau of Labor Statistics.

(Trial of Service Price Measurement in the Digital Sector)

In recent years, consumption activities using ICT devices have shifted from the purchase of digital goods themselves to the consumption of digital services. For example, images that consumers used to store and consume on their own PCs and smartphones are now stored in the cloud servers through the provision of cloud services, and the consumption of viewing services that they used to obtain by purchasing optical discs, etc.

is now increasingly consumed by purchasing streaming services through websites. Under these circumstances, insufficient quality adjustment in those services can result in an upward bias, as arose with PCs, etc. in the goods price index.

Reinsdorf and Schreyer (2017) measure the effects of digitization in computers, information and communication technology-related products, and telecommunication services, including both goods and services, and note that a quality adjustment bias has led to an upward bias in the overall price index.[55] Specifically, the price indices of 34 OECD member countries are divided into two broad categories: products clearly affected by digitalization, such as ICT equipment and telecommunications services, and products potentially affected by digitalization, such as automobiles and other consumer durables. Adding up the weight of both categories in consumer spending in 2015, Reinsdorf and Schreyer (2017) mention that they accounted for 31.5% of the total. They further multiply each weight by the bias attributable to digitization in each category and sum them, estimating an upward bias of up to +0.28 percentage points in the aggregate.

Some analyses estimate quality adjusted price indexes for cloud services using hedonic methods (Byrne, Corrado, and Sichel, 2021). In addition to the traditional characteristic variables of machine specifications (processing power, memory and storage capacity, OS type, etc.), the location of the servers is also important for cloud services. For example, for customers who require minimal latency (communication delay time), it is essential to have servers located nearby, and if they also require a backup system and supplementation (redundancy) in the event of a disaster, etc., it is desirable to have distributed servers in different locations. Using data from leading cloud service vendors, this study calculates a price index that adjusted for these characteristics for the period 2009-2014, and finds that the price decline is significant, averaging -7% per year, or in other words, that technological innovation in cloud services has been rapid. Also, Byrne, Corrado, and Sichel (2021) find that competitors' posting prices for similar services online is a trigger for further price declines. Although these approaches are mainly academic research at this time, they will be able to be reflected in the actual statistical calculations.

55 The upward bias in this study includes the quality change bias, which implies that improvements in product quality are not fully reflected in the price indexes calculated by the statistical authorities, as well as changes in utility due to the progress of digitalization, where the service price becomes close to zero (e.g., changes in demand for photography services due to the spread of smartphone's camera), and increases in utility due to the ease of access to necessary information.

(2) Adapting to New Commercial Distribution

The appearance of new commercial distribution has a variety of implications for the measurement errors. For example, e-commerce can be a source of new outlet bias because selling prices are often held down by lowering margins of wholesale and inter-dealer distribution, etc. Additionally, online platforms can use information such as online transaction and browsing history, which could lead to complex pricing decisions, including price setting for each customer. Furthermore, in contrast to consumption activities in actual stores, the purchase channel is not easily identified. Existing price survey methods may become obsolete due to these characteristics.

(Expansion of E-commerce, etc.)

In recent years, e-commerce (EC) has been expanding globally, with the consumption share via EC platforms currently around 10% in Japan and 15% in the U.S.[56] While not all goods and services sold in physical stores can be purchased online, those sold by online stores tend to be less expensive than those sold in physical stores, which can trigger the new outlet bias if there is a shift away from price-surveyed physical stores to such new shopping channels (Chart 15).[57]

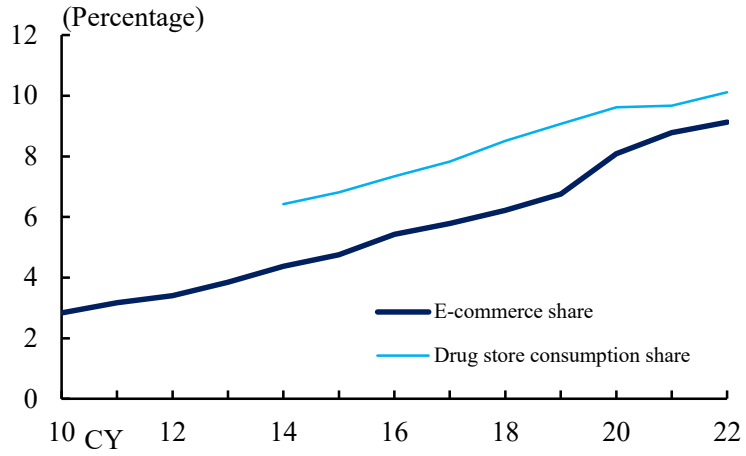
A study that directly compares the price changes of individual goods and services sold in e-commerce with the corresponding CPI items is Goolsbee and Klenow (2018), which uses price data from online retailers from 2014 to 2017 to calculate price indices for each of the items traded online, such as food and beverages, education, and clothing. They then calculate the upper level index from these price indices and compare this upper level index with the corresponding upper level index calculated from price indices in the CPI. It has been pointed out that, as a result, the price change rate calculated from online retail data is on average -1.3 percentage points lower than the CPI for each year. According to

56 The EC share here is for goods sales as reported in the "E-Commerce Market Survey" by the Ministry of Economy, Trade and Industry. The size of the BtoC-EC (e-commerce for consumers) market in Japan in 2022 was 14 trillion yen, and a breakdown of the market size in the goods sales sector shows that the largest shares were held by "Food, drinks and liquor" (2,750.5 billion yen), "Domestic electrical appliances, AV equipment, and PCs and peripherals" (2,552.8 billion yen), "Clothes and apparel goods" (2,549.9 billion yen), and "Household goods, furniture, and interiors" (2,354.1 billion yen).

57 Regarding substitution among physical stores, in Japan, the opening of large-scale stores seems to have slowed down, while drug stores are expanding their number of outlets. According to statistics from the Japan Association of Chain Drug Stores, there were approximately 10,000 stores in 2000, and 20,000 stores in 2022. Drug stores have lower prices for some daily necessities and snacks than other outlets, such as supermarkets. If these lower-priced products sold by drug stores are not surveyed, outlet bias is considered to be present.

the data used in the analysis, items traded via online shopping account for about a 19% share of the U.S. CPI. This suggests that the impact of e-commerce on the overall price index may be reasonably large, even though the share of e-commerce in U.S. retail sales is likely to be less than 20%.

Chart 15 E-commerce and Drug Stores' Share in Japan



Note: 1. The e-commerce share is the result of the E-Commerce Market Survey. The e-commerce share is the percentage of the transaction value conducted through the Internet in the household's goods consumption. Household consumption of goods is estimated based on the Family Income and Expenditure Survey and GDP statistics. The value of transactions conducted via the Internet is estimated based on a review of the literature and firm interviews, etc.

2. Drug store consumption share is an estimate using the Current Survey of Commerce, and is the share of drug stores in food sales of major retailers. Major retailers include department stores, supermarkets, convenience stores, and drug stores. The definition of drug stores is based on the Japan Standard Industrial Classification.

Source: Ministry of Economy, Trade and Industry.

(Diversification in Pricing)

Another aspect of the new commercial distribution is the increasing complexity and sophistication of pricing. For example, the appearance of dynamic pricing with real-time demand monitoring, the expansion of subscription services and the further use of two-part tariffs (a type of non-linear pricing), the spread of point systems in e-commerce, and the possibility that the provision of search history data, etc., may have begun to amount to a transfer of value as the use of data becomes more valuable (a type of barter through

data, rather than money).[58] See Appendix 7 for more on these topics. The discussion of dynamic pricing, for example, has some affinity with the idea of the Dynamic Price Index proposed by Reis (2005),[59] and the point system offered by online marketplaces can be categorized as a straightforward upward bias, except for the measurement problem of the monetary value of points. However, the accumulation of research on how the COLI concept should be reconstructed under these new commercial trends has not yet progressed significantly.

6. Conclusion

When central banks of major economies set numerical targets for price stability, the CPI is widely used. In addition, the measurement errors (upward bias) of the CPI are often cited as one of the reasons for setting the numerical targets for price stability at a positive value. Therefore, this paper reviewed recent research trends on the type and magnitude of measurement errors in the CPI, with reference to cases in overseas countries as well as Japan.

More than 25 years have passed since the publication of the 1996 Boskin Report, which stated that there was an upward bias of +1.1 percentage points in the U.S. CPI. The bias derived from the factors identified in the Boskin Report seems to be decreasing on the whole, not only in the U.S. but also in Japan and Europe, although there are differences between countries and among bias types. This decrease is mainly due to the fact that the statistical agencies in each country have continuously worked to improve the accuracy of their indices, while being aware of the existence of measurement errors, for example by

58 Using Korean retail data, Ueda, Watanabe and Watanabe (2022) compare the differences between online and offline price trends, as well as trends in listed prices and actual transaction prices. The study reports that online prices are lower and more elastic than offline prices for the same product sold by the same firm, and that online transaction prices vary widely, reflecting customer-specific discounting factors. When such price variation exists due to differences among purchasing channels and customer factors, it is not easy to select representative transactions for price surveys.

59 The COLI discussed in Appendix 2 is derived by assuming that all given income earned in the current period is spent on goods and services in the current period (in other words, eliminating savings behavior). In fact, however, consumers are likely to change the combination of goods and services at different points in time; for example, if today's price is higher than tomorrow's, they will shift their spending to tomorrow. The pioneering work by Alchian and Klein (1973) argues that the price indices should include asset prices once households' intertemporal decision-making regarding expenditure is considered, as asset prices contain information about "prices of present claims on future consumption." Reis (2005), extending the work of Alchian and Klein (1973) to the setting of incomplete insurance markets, theoretically derives the dynamic price index (DPI). He argues that DPI includes asset prices and distinguishes between durable and non-durable goods' prices. For other DPI studies, see, for example, Aoki and Kitahara (2010).

increasing the frequency of updating the expenditure weights used to aggregate the item indices into upper level indices, by changing the formulas used to calculate the indices, and by extending the quality adjustment methods to cover a wide range of items. Changes in the structure of the economy may also have contributed to the reduction of the measurement errors: the decline in the price of ICT equipment, which was pronounced at the time of the Boskin Report, has slowed down in recent years, and the substitution of private retailers by GMS has stopped.

The trend of declining measurement errors does not necessarily mean that they have now disappeared and will not increase again in the future. For instance, there could be residual measurement errors in the prices of some goods and services in the CPI due to insufficient identification of the contribution of quality differences to observed price changes. The reasons for this insufficiency include the fact that the method for identifying portions of quality change and eliminating the quality change bias were not academically established in the first place, or there are practical limitations such as lack of data and manpower constraints when conducting quantitative evaluations. It has also been pointed out that changes in commercial distribution, such as the recent expansion of e-commerce, which were not the focus of attention at the time of the Boskin Report's release, could have affected the bias. Research and analysis on this issue has only just begun. Furthermore, if the economic structure changes more in the future, such as the growth of the service sector, the digitalization of the economy, and the aging of the population, it may become more difficult to apply existing price evaluation methods, and the magnitude of measurement errors could be affected. Given that measurement errors cannot be eliminated in this way, the measurement errors of the CPI continue to have validity as one of the reasons for setting numerical targets for the CPI.

Appendix 1. Functional Form Identification of Price Indexes Based on the Axiomatic Approach

One approach to identifying index formulas for price statistics is the axiomatic approach. The axiomatic approach derives the functional form from such axioms, although the desired properties (axioms) of the price index differ depending on the way price statistics are used. For example, there is a property called "commensurability," which means that the index value does not change when the unit of measurement for prices is changed. This property implies that the price index remains the same whether the price of gasoline is measured in liters or gallons, and is one of the properties that price indexes are expected to satisfy. In this manner, the axiomatic approach derives the functional form axiomatically or deductively, based on the properties required of the exponent as a starting point. In Appendix 1, we review the axiomatic approach with reference to Abe (2023).

In general, monotonicity, linear homogeneity, identity, commensurability, and dimensionality are often pointed out as the basic properties required of price indexes (Chart A-1). Monotonicity refers to the property that the price index is an increasing function with respect to the comparison period price and a decreasing function with respect to the base period price, requiring the overall price index to increase when each item price increases in the time series. Linear homogeneity is the property that when a constant multiple of a price vector (each item price) is taken, the upper-level price index that aggregates these item prices is also a constant multiple. Identity refers to the property that if the price does not change, the index will not change. Commensurability, as noted above, requires that the value of the index be the same regardless of which system of units is used to measure prices. Dimensionality is the property that a change in the currency unit does not affect the price index.

There are several other axioms that are also considered important (Chart A-2). For example, transitivity is the property that allows the time path of prices to be decomposed. When this property is present, the price index between two points in time can be expressed as the product of the two price indexes before and after one point in time between the two start and end points. Circularity is an extension of transitivity and refers to the property that the level of the price index returns to its original level after one round of the time path regarding price changes.

Chart A-1 Axioms of Price Indexes (1): Basic Properties

✓ Monotonicity

Price index is increasing (decreasing) function of the comparison-period price p_1 (base-period price p_0):

$$PI(\boldsymbol{q_0}, \boldsymbol{p_0}, \boldsymbol{q_1}, \boldsymbol{p_1}) > PI(\boldsymbol{q_0}, \boldsymbol{p_0}, \boldsymbol{q_1}, \boldsymbol{p}) \quad \text{if } \boldsymbol{p_1} \geq \boldsymbol{p}$$
$$PI(\boldsymbol{q_0}, \boldsymbol{p_0}, \boldsymbol{q_1}, \boldsymbol{p_1}) < PI(\boldsymbol{q_0}, \boldsymbol{p}, \boldsymbol{q_1}, \boldsymbol{p_1}) \quad \text{if } \boldsymbol{p_0} \geq \boldsymbol{p}$$

✓ Linear homogeneity

The constant multiple of the price vector makes the price index the same constant multiple:

$$PI(\boldsymbol{q_0}, \boldsymbol{p_0}, \boldsymbol{q_1}, \alpha \boldsymbol{p_1}) = \alpha PI(\boldsymbol{q_0}, \boldsymbol{p_0}, \boldsymbol{q_1}, \boldsymbol{p_1})$$

✓ Identity

If the price does not change, the index will be 1:

$$PI(\boldsymbol{q_0}, \boldsymbol{p_0}, \boldsymbol{q_1}, \boldsymbol{p_1}) = 1 \quad \text{if } \boldsymbol{p_0} = \boldsymbol{p_1}$$

✓ Commensurability

The index value is the same no matter which system of units is used to measure price. Converting 1 liter of water to 1,000 milliliters and converting the price from 100 yen/liter to 0.1 yen/milliliter does not change the price index.

✓ Dimensionality

A change in the currency unit has no effect on prices. Major indices satisfy this property:

$$PI(\boldsymbol{q_0}, \alpha \boldsymbol{p_0}, \boldsymbol{q_1}, \alpha \boldsymbol{p_1}) = PI(\boldsymbol{q_0}, \boldsymbol{p_0}, \boldsymbol{q_1}, \boldsymbol{p_1})$$

Chart A-2 Axioms of Price Indexes (2): Other Properties

✓ Transitivity

Price time paths can be decomposed:

$$P(s,t) \cdot P(t,k) = P(s,k)$$

✓ Circularity

When the price time path completes one cycle, the index also reverts to its original value:

$$P(s,t) \cdot P(t,s) = 1$$

Under the axiomatic approach, it is shown that the only price index that satisfies the five main axioms mentioned above is a Cobb-Douglas type function (Funke, Hacker, and Voeller, 1979):

$$P(0,\ 1) = \prod_{i=1}^{N} \left(\frac{p_{i,1}}{p_{i,0}}\right)^{\alpha_i}, \qquad \sum_{i=1}^{N} \alpha_i = 1 \qquad (1)$$

While the Cobb-Douglas type has desirable axiomatic properties, no objective criteria are given for selecting the weights α_i. In other words, it does not have answers to basic and important issues such as whether the CPI weight should be based on a percentage of expenditures at a single point in time, and whether the weight should be calculated at the base or comparison period. Thus, when a price index attempts to satisfy many properties, the functional form is severely constrained. Therefore, it is essential to determine which axioms should be satisfied and which can be discarded, depending on the purpose of the price index.

Appendix 2. The Economic Approach and the Superlative Index

In Appendix 2, we review the identification of price indices using the economic approach and the position of the superlative indexes under this approach, with reference to Abe (2023) and Shiratsuka (1998).[60] The economic approach focuses on affinity with economic theory (consumer theory). Specifically, the step of this approach is to take the utility function as a starting point and define the price index as the two-point ratio of the expenditure function necessary to maximize household utility and maintain that utility level. This calculation method is that of deriving COLI as outlined in Section 2. In such cases, general utility and expenditure functions should be used, and in view of the practical treatment, it is preferable not to require the estimation of price elasticity of demand and cross-elasticity. When calculating COLI, the homothetic utility functions are often chosen, for example, the Cobb-Douglas type or constant elasticity of substitution (CES). This choice is made because the COLI function is easier to handle in cases where COLI is determined independently of each household's income level.

60 Abe (2023) has more details on the economic approach. For a detailed mathematical and economic definition of the superlative index and its properties, see Diewert (1976, 1978) and Yoshizoe (2017). The main purpose of this section is to provide an overview of the previous studies on the relationship between COLI and the superlative index, setting aside mathematical exactness to some extent.

In this paper, COLI is derived based on a CES-type utility function as a practical example:

$$U = \left\{ \sum_{i=1}^{N} (a_i \, q_i)^{\frac{\sigma-1}{\sigma}} \right\}^{\frac{\sigma}{\sigma-1}} \qquad where \;\; \sigma > 1, \;\; a_i > 0, \;\; \sum_{i=1}^{N} a_i = 1 \qquad (2)$$

where N is the total number of goods, q is consumption volume, and σ is the elasticity of substitution. Solving the utility maximization problem under the budget constraint equation,

$$\max U(\boldsymbol{q}) \quad s.t. \;\; \sum_{i=1}^{N} p_i \, q_i = Y \qquad (3)$$

the expenditure function is derived ($\boldsymbol{q}$ and $\boldsymbol{p}$ are vectors whose components are the quantity and price of consumption of each good, and Y is nominal gross income):

$$C(\boldsymbol{p}, u) = Y = \left\{ \sum_{i=1}^{N} \left(\frac{p_i}{a_i} \right)^{1-\sigma} \right\}^{\frac{1}{1-\sigma}} \cdot u \qquad (4)$$

Since the ratio of the expenditure function between two points in time that achieve the same utility is equal to the COLI, the COLI function is derived as follows:

$$PI(\boldsymbol{p_0}, \boldsymbol{p_1}) = \frac{C(\boldsymbol{p_1}, \, u)}{C(\boldsymbol{p_0}, \, u)} = \left\{ \frac{\sum_{i=1}^{N} a_i^{\sigma-1} \, p_{i1}^{1-\sigma}}{\sum_{i=1}^{N} a_i^{\sigma-1} \, p_{i0}^{1-\sigma}} \right\}^{\frac{1}{1-\sigma}} \qquad (5)$$

where $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$ represent the price vectors at the base and comparison periods. Equation (5) is the COLI derived from the CES-type utility function. However, because this form requires measurement of elasticity of substitution, which is difficult to handle in practice, several derivative forms exist, such as the Sato-Vartia index (Sato, 1974, 1976; Vartia, 1976), which does not require elasticity. These price indexes are said to be "exact index," which is consistent with the COLI under certain conditions, with the CES-type utility function.

To be implemented in statistical practice, the price index calculating method must be able to withstand timely publication. As mentioned above, when exact indices are used

as price indices for CES-type utility functions, the elasticity of substitution must be estimated, except in the Sato-Vartia index. In these cases, there are issues such as how to handle this estimation in practice, the reliability of the estimation accuracy, and its impact on measurement errors. Therefore, the "superlative index" was proposed based on the idea of obtaining an approximation of the exact price index against the COLI derived from a specific utility function or expenditure function, without the parameter estimation.

To derive the superlative index, we assume a general and flexible utility function as a starting point. This approach is to find a function that is a good approximation of a more flexible utility function and adopt this function as the price index. Specifically, the approach results in the problem of finding a second-order approximation function f to f^*, which are second-order differentiable functions that map mathematically in the N-dimensional positive-valued real number space:

$$\begin{aligned} f(\boldsymbol{x}^*) &= f^*(\boldsymbol{x}^*) \\ \nabla f(\boldsymbol{x}^*) &= \nabla f^*(\boldsymbol{x}^*) \\ \nabla^2 f(\boldsymbol{x}^*) &= \nabla^2 f^*(\boldsymbol{x}^*) \end{aligned} \qquad (6)$$

Assuming first-order homogeneity, which is a standard assumption in economics, the function known as the "quadratic mean of order r" has been shown to be consistent with "a second-order approximation of any second-order continuously differentiable function" (Diewert, 1976).

$$f(\boldsymbol{p}) = \left(\sum_{i=1}^{N} \sum_{k=1}^{N} a_{ik}\, p_i^{r/2}\, p_k^{r/2} \right)^{\frac{1}{r}} \qquad (7)$$

$$where \quad a_{ij} = a_{ji},\ r \neq 0$$

This property is extremely powerful, and the price index that corresponds to this flexible function is regarded as the superlative index. By using this superlative index, the COLI can be obtained as a good approximation of an index that is exact for a certain utility and expenditure function, without assuming a specific utility and expenditure function. In this case, parameter estimation is not required. This function corresponds to the Walsh index when $r = 1$ and the Fisher index when $r = 2$, and both indices are defined as the superlative index.

The trans-log type function is also classified as a flexible function other than the quadratic mean of order r. Thus, the trans-log type function also satisfies the condition of equation (6). We therefore assume a trans-log type expenditure function:

$$\ln C(\boldsymbol{p}, u) = a_0 + \sum_{i=1}^{N} a_i \ln p_i + \frac{1}{2} \sum_{i=1}^{N}\sum_{k=1}^{N} a_{ik} \ln p_i \ln p_k + b_0 \ln u + \sum_{i=1}^{N} b_i \ln p_i \ln u + \frac{1}{2} b_{00} (\ln u)^2 \quad (8)$$

$$where \qquad a_{ik} = a_{ki}, \ \sum_{i=1}^{N} a_i = 1, \ \sum_{k=1}^{N} a_{ik} = 0, \ \sum_{i=1}^{N} b_i = 0$$

Calculating the COLI based on this expenditure function, the Törnqvist index (in logarithmic form) is derived:

$$\ln PI(\boldsymbol{p_0}, \boldsymbol{p_1}) = \ln C(\boldsymbol{p_1}, \sqrt{u_0\, u_1}) - \ln C(\boldsymbol{p_0}, \sqrt{u_0\, u_1}) = \frac{1}{2}\sum_{i=1}^{N} (w_{i0} + w_{i1})(\ln p_{i1} - \ln p_{i0}) \quad (9)$$

As mentioned above, since the trans-log type function is flexible, by definition, the Törnqvist index is also classified as the superlative index. Incidentally, the trans-log type function corresponds to a special form of Quadratic Mean of Order r ($r = 0$) with the additional assumption (Diewert, 1980). The trans-log expenditure function and the trans-log utility function are not dual, and there is no factor reversal between the Törnqvist price index and the Törnqvist quantity index. It should be added, however, that quantity and price indexes do not necessarily have to be figured in the same functional form. Not only does the Törnqvist index satisfy the properties of the superlative index, but the chained Törnqvist index is also regarded as a discrete-time approximation of the Divisia index (an index defined in continuous time) that satisfies many axioms. Therefore, the Törnqvist index is widely used. However, unlike the Divisia index, the chained Törnqvist index is not always universal, since the chained Törnqvist index does not have transitivity and has the problem of "drift phenomenon" (see the footnote below). Since this topic is

beyond the scope of this paper, we will not discuss it in detail, but this fact can pose a serious problem when calculating price indexes using high-frequency data such as scanner data.

Appendix 3. Functional Form of Major Price Indices

Chart A-3 summarizes the most commonly used indices when aggregating item indices into upper-level indices. The Laspeyres index is used in the formal series of price statistics and is an arithmetic mean with weights at the base period, whereas the Paasche index uses comparison period information for the weights (Drobisch, 1871; Laspeyres, 1871; Paasche, 1874). The Fisher index (Fisher, 1922) is the geometric mean of the Laspeyres and Paasche indices, and the above three indices are widely known due to the similarity of their calculation methods. As mentioned above, the Törnqvist index is the geometric mean using the average weights at the base and comparison periods, and the Walsh index is the arithmetic mean using the geometric mean weights at the base and comparison periods. The Fischer, Törnqvist, and Walsh indexes are considered the superlative indexes, as mentioned in Appendix 2.

Chart A-4 summarizes the lower-level aggregation method, i.e., the index formula for aggregating brand/product prices into an item index. Roughly categorized, there are the arithmetic and geometric averages of the price levels of surveyed brands/products within an item, or those averages of the price indexes by surveyed brands/products. The Dutot index (Dutot, 1738) corresponds to a ratio of the average price of all brands/products at two points in time. The Jevons index (Jevons, 1863) is calculated as a simple geometric mean of the price ratios for each brand/product with respect to the base period. In the Carli index (Carli, 1764), the index value is derived by calculating a simple arithmetic mean of the price ratios for each brand/product with respect to the base period. In general, within-item aggregation should be determined according to the axiomatic properties required by each country (such as transitivity and commensurability) and the quality/price variance among the brands/products. The European HICP standard can use either the Dutot or Jevons index, while the Carli or Jevons index is used in the U.S. This formula difference could be due to differences in the variance of quality and price within items caused by the difference in the number of items between the two areas. If the Dutot index were used for within-item aggregation in the U.S., it could be assumed that the weights of expensive brands/products would be larger, which means that the bias can exist when the price representativeness of these brands is low.

Chart A-3 Formula Used to Aggregate the Item Indexes to the Upper-level Index

Laspeyres index: arithmetic average with base period weights

$$PI^{L}(\boldsymbol{p_0}, \boldsymbol{p_t}) = \frac{\sum_{i=1}^{N} p_{it}\, q_{i0}}{\sum_{i=1}^{N} p_{i0}\, q_{i0}} = \sum_{i=1}^{N} w_{i0} \left(\frac{p_{it}}{p_{i0}}\right)$$

Paasche index: arithmetic average with comparison period weights

$$PI^{P}(\boldsymbol{p_0}, \boldsymbol{p_t}) = \frac{\sum_{i=1}^{N} p_{it}\, q_{it}}{\sum_{i=1}^{N} p_{i0}\, q_{it}}$$

Fisher index: geometric mean of Laspeyres and Paasche index

$$PI^{F}(\boldsymbol{p_0}, \boldsymbol{p_t}) = \sqrt{PI^{L} \cdot PI^{P}}$$

Törnqvist index: geometric mean using average weights at base and comparison periods

$$PI^{T}(\boldsymbol{p_0}, \boldsymbol{p_t}) = \prod_{i=1}^{N} \left(\frac{p_{it}}{p_{i0}}\right)^{\frac{w_{i0}+w_{it}}{2}}$$

Walsh index: arithmetic mean using geometric mean weights at base and comparison periods

$$PI^{W}(\boldsymbol{p_0}, \boldsymbol{p_t}) = \frac{\sum_{i=1}^{N} p_{it}\, \sqrt{q_{it} q_{i0}}}{\sum_{i=1}^{N} p_{i0} \sqrt{q_{it} q_{i0}}}$$

Chart A-4 Formula Used to Aggregate the Brand/Product Prices or Indexes within an Item

Dutot index: ratio of average price between two points in time

$$PI^{D}(\boldsymbol{p_0}, \boldsymbol{p_t}) = \frac{\sum_{i=1}^{N} p_{it}}{\sum_{i=1}^{N} p_{i0}} = \frac{\frac{1}{N}\sum_{i=1}^{N} p_{it}}{\frac{1}{N}\sum_{i=1}^{N} p_{i0}}$$

Jevons index: simple geometric mean of price ratio from base period

$$PI^{J}(\boldsymbol{p_0}, \boldsymbol{p_t}) = \prod_{i=1}^{N} \left(\frac{p_{it}}{p_{i0}}\right)^{\frac{1}{N}}$$

Carli index: simple arithmetic mean of price ratio from base period

$$PI^{C}(\boldsymbol{p_0}, \boldsymbol{p_t}) = \frac{1}{N} \sum_{i=1}^{N} \frac{p_{it}}{p_{i0}}$$

Appendix 4. Fundamental Information on Price Indexes for Each Country

Chart A-5 summarizes fundamental information on price indices in Japan, the U.S., and Europe. Each country has its own concept of price indexes and calculating methods. The causes of these differences can include differences in the number of firms in each item and market in respective countries, differences in the number of products and services to be surveyed, and differences in the geographic dispersion of the stores to be surveyed. Regarding the concept of the price index, in the U.S., the CPI is positioned as the COLI, while the CPI in Japan and Europe is not positioned as the COLI. In terms of the survey method, for example, the U.S. CPI has only about half as many items as the Japanese and European CPIs, and the number of survey prices is also smaller. This discrepancy is related to the fact that the U.S. uses a sampling with probability proportional to size, while Japan and Europe use the method that surveys representative stores and brands/products. As a result of the large number of items, the scope of each item is limited in Japan and Europe, and the quality of the surveyed brands/products within each item is homogeneous. Also, the number of survey prices is higher than the U.S. In the U.S., where the number of items is small, the item index is constructed by sampling with probability proportional to expenditure amount, given the variation in surveyed brands/products within an item.

Chart A-5 Fundamental Information on CPI by Country[1]

	Japan	The U.S.	Europe (HICP/ National standard)
Statistical positioning	Not COLI[2]	**COLI**	Not COLI (HICP standard)[3]
Number of items	582	**211**	About 650 (DE)/ 410 (FR)/ 422 (IT)
Number of surveyed stores	About 28,000	About 23,000	About 30,000 (DE/ FR)/ About43,000 (IT)
Number of surveyed prices	About 210,000[4]	**About 94,000**	About 350,000 (DE) About 200,000 (FR) About 332,000 (IT)
Store and price sampling method	Representative stores and brands/ products	**Sampling with probability proportional to size**	Representative stores and brands/ products
Within-item formula[5]	Dutot index	Jevons index (61%)/ Carli index (39%)[6]	Dutot index (DE)/ Jevons index (FR/ IT)[7]
Inter-item formula	Fixed-weight base Laspeyres index	Fixed-weight base Laspeyres index	Chained Laspeyres index (HICP standard/ FR/ IT) / Fixed-weight base Laspeyres index (DE)
Weight change interval	5 years (with Midpoint-year Revision)	2 years[8]	1 year (HICP standard/ FR/ IT)/ 5 years (DE)
New product review intervals	2-3 years "Midpoint-year Revision"	1 year	1 year (HICP standard/ FR/ IT) / 5 years in principle (DE)

Note: 1. Based on various references and published information from the Ministry of Internal Affairs and Communications, Bureau of Labor Statistics, and eurostat. Some of the figures may therefore not be up-to-date, and the time points for each category could be slightly different. DE=Germany, FR=France, and IT=Italy.

2. On the website, it is mentioned that the CPI does not measure changes in the cost of living associated with changes in the type, quality, or quantity of goods purchased due to changes in household lifestyles or preferences, etc.

3. The website states, "The HICP is designed to assess price stability and is not intended to be a cost of living index."

4. For on-site survey. In addition, approximately 6 million prices are surveyed, including approximately 750,000 prices for POS data survey (PCs, etc.) and approximately 5 million prices for web scraping.

5. The Dutot index is "the price ratio of the arithmetic mean of the prices of each brand/product at the point of comparison and base period (the ratio of the arithmetic mean price at each point in time)," the Jevons index is "the simple geometric mean of the price ratio," and the Carli index is "the simple arithmetic mean of the price ratio."

6. Carli index before 1998. 7. The HICP standard permits both formulas. 8. The weights are updated every year starting in 2023.

Sources: Ministry of Internal Affairs and Communications; Bureau of Labor Statistics; eurostat; Umeda (2009); ECB (2021); Collin (2006); Ministry of Internal Affairs and Communications (2019).

Appendix 5. Major Actions to Improve Price Index Accuracy

Since the publication of the Boskin Report, statistical agencies in each country have continued their efforts to improve price index accuracy (Chart A-6). To cite an example of an improvement solution, the U.S. introduced geometric averaging within some items in 1999 to address the lower-level substitution bias. In order to reduce quality change bias, statistical agencies in various countries are also progressively expanding the coverage of hedonic methods. The use of hedonic methods in the U.S. and Europe was in place even before the Boskin Report. In Japan, the CPI revision to the 2000-base was taken as an opportunity to apply them simultaneously to IT products. Regarding weight changes, the U.S. significantly reduced the update interval in 2002 (to be updated annually from 2023, as in Europe), the European HICP standard made annual updates a rule in 2012, and Japan introduced the Midpoint-year Revision at the time of the transition to the 2000-base. From this table, it can be seen that the situation surrounding the measurement errors of the CPI in Japan has changed significantly since the 2000-base.

Chart A-6 Major Efforts to Improve Index Accuracy

	Japan	U.S.	Europe (HICP/ National standard)
1980s		Application of hedonic method to housing (1988)	Application of hedonic method to books (1987, FR)
Early 1990s		Application of hedonic method to apparel (1991)	
Late 1990s		✓ **Publication of the Boskin Report** (1996) ✓ **Formal application of geometric averages** within some items (1999) ✓ Application of hedonic method to televisions (1999)	Application of hedonic method to dishwashers (1997, FR)
2000		Application of hedonic method to college textbooks, white goods, and video equipment	Application of hedonic method to men's long sleeve shirts (FR)
2001 (Release of 2000-base CPI)	✓ Since 2000-base, application of hedonic method to IT devices started. ✓ Since 2000-base, **Midpoint-year Revision** system has been introduced to evaluate whether or not items should be revised or eliminated between base years (every 5 years).		
2002		**Shorten weight change interval from 10 to 2 years**	Application of hedonic method to PC (DE)
2003	Revised the method of setting price survey districts (division of districts according to topographic features)		
2012			**Rule for annual update of weights** (HICP standard)
2019		Application of hedonic method to fixed telephone service, Internet services, and cable & satellite television service	
2021			Proposed inclusion of OOH cost in HICP
2023		**Shorten weight change interval from 2 to 1 year (annually)**	

Note: Based on various literature and published information from the Ministry of Internal Affairs and Communications and the Bureau of Labor Statistics. FR=France and DE=Germany.

Sources: Ministry of Internal Affairs and Communications; Bureau of Labor Statistics; ECB (2021); Umeda (2009).

Appendix 6. Upper-level Substitution Bias by Country

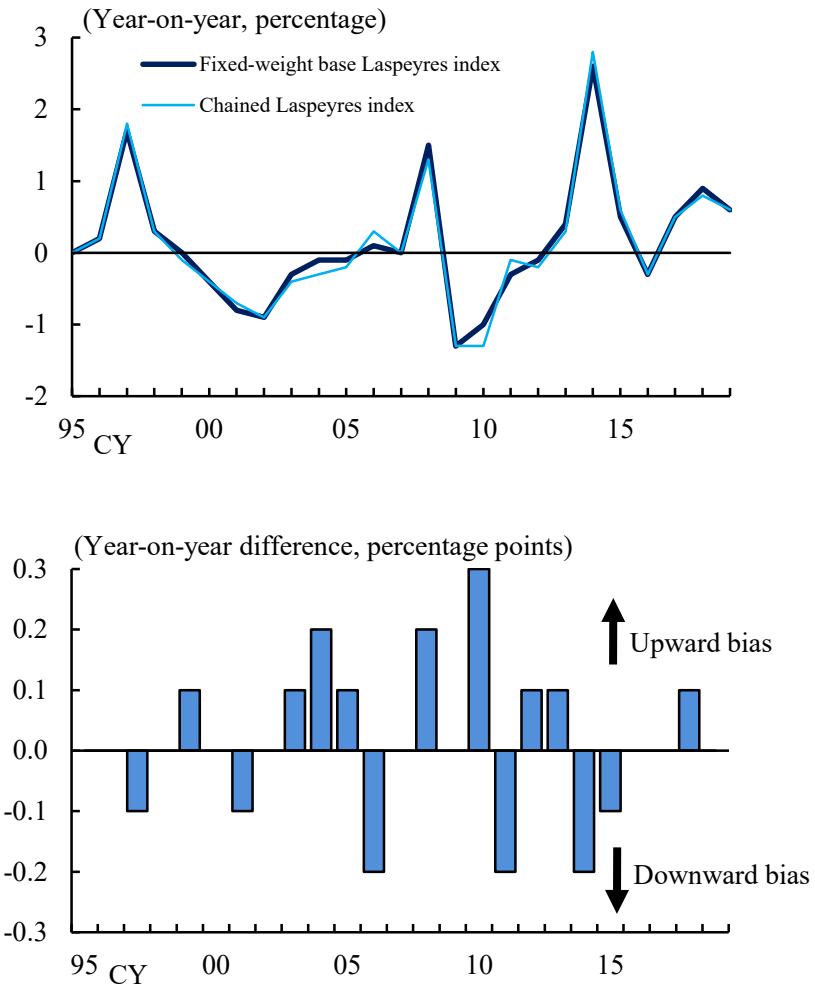
In many previous studies, the estimation of the magnitude of the upper-level substitution bias is conducted by calculating the difference between the superlative index, the chained Törnqvist index (or the chained Laspeyres/Paasche index, which are published as reference indices), and the main index of the CPI. For example, in the case of the chained Törnqvist Index, the weights used to aggregate item indexes into the upper-level index are the average of the weights at the base and comparison period, and the aggregation formula is the geometric mean, which is similar to the demand function when a Cobb-Douglas type utility function is assumed (the demand function corresponding to the Cobb-Douglas type utility function is the geometric mean of each item index aggregated with the expenditure weights at each point in time). Therefore, it is not unnatural to consider the difference between the main CPI index and the chained Törnqvist index as an estimated measurement bias.[61, 62] In this appendix, following this approach, we summarize the magnitude and characteristics of the upper-level substitution bias in Japan, Europe, and the U.S.

Chart A-7 shows the year-on-year change in the fixed-weight base Laspeyres index (the main index of CPI) and the chained Laspeyres index, as well as the difference between the two for Japan. Although there were phases in which the upper-level substitution bias expanded until the 2000s, it has generally remained near zero since the 2010s, when the price decline of ICT equipment slowed down. Considering these trends, it can be assessed that Japan's upper-level substitution bias is currently at a low level of at most +0.1 percentage points.

61 It should be noted that the bias defined by the difference between the fixed-weight base and chained Laspeyres indexes can include not only the narrowly defined upper-level substitution bias (demand substitution among items), but also the effects due to the drift phenomenon and reset effect, and technical effects by formula (such as the presence or absence of transitivity). For example, in the case of a chained index, when there are items with price bouncing, there is a drift phenomenon where the index level does not return to its original level even if the price level returns. In addition, because of the characteristic of the fixed-weight base Laspeyres index of having different contributions to the overall index depending on the index level of each item, when the index base is revised to reset the index level back to the original level, there will be a difference between the year-on-year rate of the old and new indexes (reset effect).

62 However, it is difficult to publish the chained Törnqvist index as the main index from the perspective of rapidity. Even in the U.S., there is a lag of several years before information on actual expenditure weights for a given month is finally reflected in the statistics.

Chart A-7 Upper-level Substitution Bias of the Japanese CPI

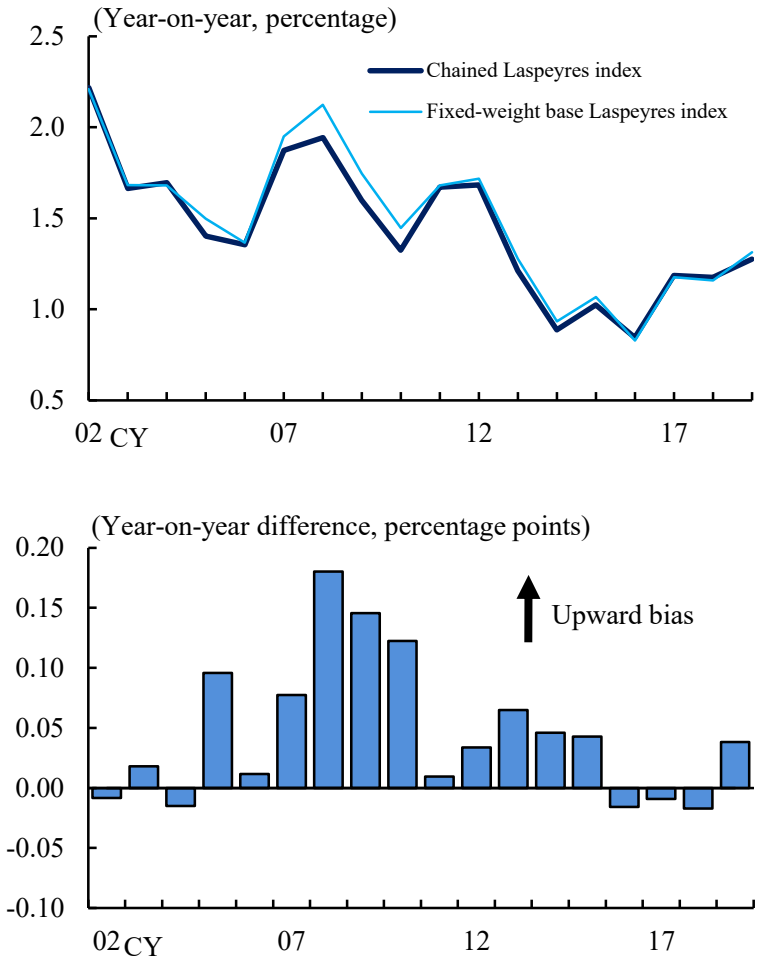


Note: CPI (all items less fresh food) including consumption tax effects. The lower chart shows the year-to-year deviation of the fixed-weight base Laspeyres index (the main index) from the chained Laspeyres index.

Source: Ministry of Internal Affairs and Communications.

In Europe, the HICP basket, expenditure weights, has been changed to be updated annually in 2012. Therefore, there is no need to assess the upper-level substitution bias for the series after this year. Looking at the fixed-weight base Laspeyres index and chained Laspeyres index since 2002 in Chart A-8, it is clear that the upper-level substitution bias remained low, around +0.1 percentage points, even before 2011, when the weights were not updated annually.

Chart A-8 Upper-level Substitution Bias of the CPI in Europe

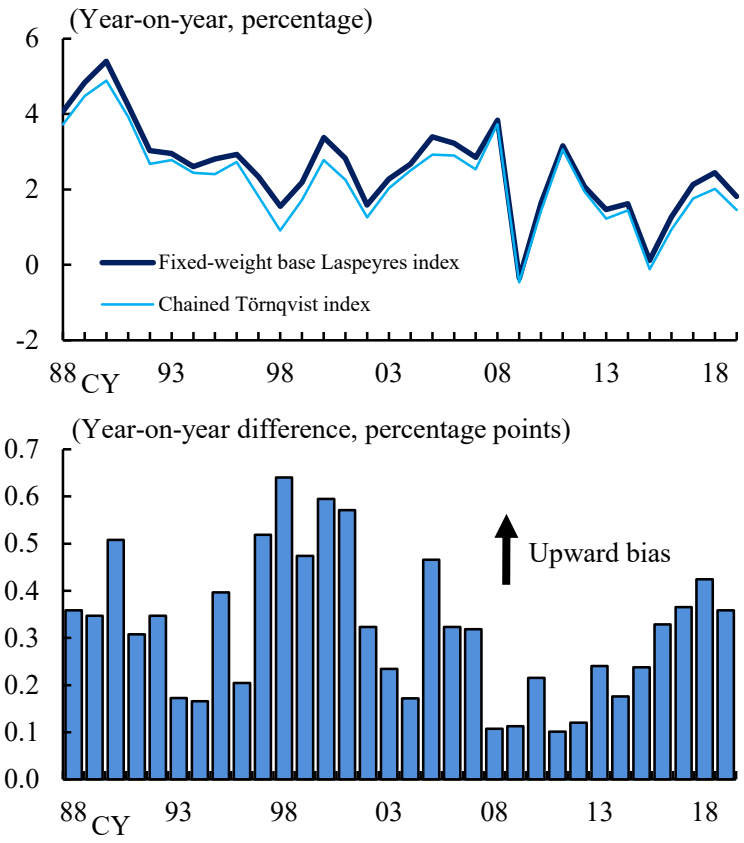


Note: HICP (all items excluding energy, food, alcohol and tobacco). The fixed-weight base Laspeyres index is hand-calculated estimates based on weighted average of the medium classification indexes (87 series) with the base years 2000, 2005, 2010, and 2015. The lower chart shows the year-to-year deviation of the fixed-weight base Laspeyres index from the chained Laspeyres index.

Source: eurostat.

Chart A-9 shows, for the U.S., the year-to-year change in the fixed-weight base Laspeyres index (the main index of CPI) and the chained Törnqvist index, as well as the difference between the two. In the mid-1990s, just before the Boskin Report was published, the year-to-year deviation was about +0.2 percentage points. Subsequently, from the late 1990s to the early 2000s, it rose to over +0.5 percentage points. Furthermore, the deviation widened in the mid-2000s and late 2010s, and the bias has not been continuously decreasing.[63] Also, the current level is +0.3 percentage points higher than in Japan and Europe.

Chart A-9 Upper-level Substitution Bias of the U.S. CPI



Note: CPI (all items). The lower chart shows the year-to-year deviation of the fixed-weight base Laspeyres index (the main index) from the chained Törnqvist index. The pre-2000 chained Törnqvist index is an estimate by Cage, Greenlees, and Jackman (2003). Data since 2001 are published by the BLS.

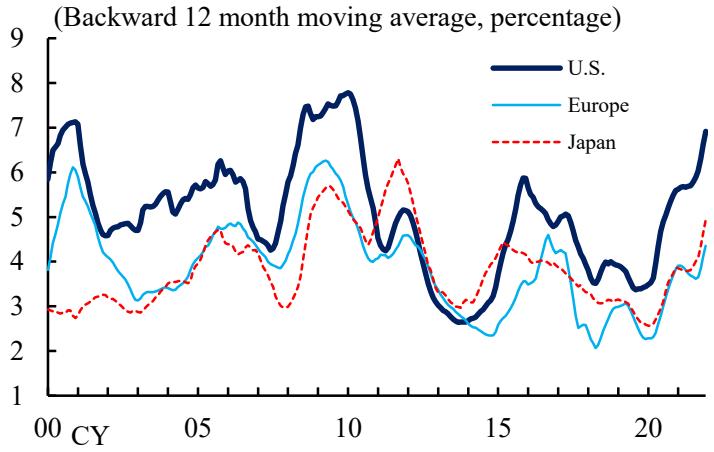
Sources: Bureau of Labor Statistics; Cage, Greenlees, and Jackman (2003).

63 Gordon (2006) found that the upper-level substitution bias, as measured by the deviation between the chained Törnqvist index and the main index, averaged +0.38 percentage points from January 2000 to January 2006, which is larger than the +0.15 percentage points in the Boskin Report. Also, given that the BLS had implemented a number of solutions to improve index accuracy by 2006, such as increasing the frequency of weighting updates, the study noted that the value of measurement errors in the Boskin Report could have been underestimated.

Since the upper-level substitution bias is due to shifts in demand triggered by changes in the relative price of individual goods or services, its magnitude and variability can be related to the dispersion of relative price differentials. Based on this perspective, Lebow and Rudd (2003) noted that the widening of the bias in the U.S. in the late 1990s might be due to the large dispersion of relative price changes. Cage, Greenlees, and Jackman (2003) also pointed out that the large price decline in "personal computers and peripheral equipment" and the large price increase in "utility natural gas services" between 1999 and 2000 may have contributed to the bias widening.[64]

Such variation in relative price fluctuations could explain the differences in the magnitude of the upper-level substitution bias between Japan, the U.S., and Europe. Chart A-10 shows the variation in the change rate of the CPI at the subgroup level. The level of the U.S. is permanently higher than that of Japan and Europe, and its value also fluctuates widely up and down.[65]

Chart A-10 Variation in Price Change Rate by Subgroup



Note: Standard deviation of the CPI year-on-year change rate by subgroup (EU: HICP-based). The number of subgroup indexes in December 2021 are as follows: U.S.: 69 series; EU: 86 series; Japan: 104 series. Japan includes the impact of the consumption tax.

Sources: Ministry of Internal Affairs and Communications; Bureau of Labor Statistics; eurostat.

64 Cage, Greenlees, and Jackman (2003) used data from December 1999 to December 2000 to calculate and compare the Laspeyres and chained Törnqvist indices on a basis that excludes from the aggregate the items with large relative price fluctuations. The paper reported that excluding items with high fluctuation results in a smaller divergence between the two indices, compared with the non-excluded case.

65 A similar trend can be observed when the variation indices are calculated and compared for Japan and the U.S. at a more detailed item level.

Appendix 7. Increasingly Diverse and Complex Pricing and Price Surveys

Dynamic pricing is often applied to online transactions, such as hotel charges and airplane fares. In general, prices are adjusted in response to supply and demand based on real-time aggregation of Internet bookings. Web scraping technology, which extracts pricing information from websites, can be effective in capturing dynamic pricing from the perspective of price surveys by statistical authorities. Japan's CPI has started using web-scraping technology in its price surveys since the 2020-base, and introduced it to price surveys for hotel charges, charges for overseas package tours, and airplane fares. In dynamic pricing, prices may fluctuate significantly when reservations are made shortly before the date of use, reflecting reservation status and other factors. Therefore, a rule has been adopted for hotel charges and charges for overseas package tours that do not include prices for last-minute bookings, as well as noise treatment, such as eliminating extreme movements. Similar efforts to use the technology are ongoing in other countries (Chart A-11).

Chart A-11 Status of Web Scraping Technology Use by Country

	Applied items	Items under study and review
Japan	✓ Hotel charges ✓ Charges for overseas package tours ✓ Airplane fares	Price survey via online supermarket website information: ✓ Food ✓ Medicine ✓ Daily necessities
Netherlands	✓ Apparel	✓ Airplane fares
Italy	✓ Train ticket	✓ Airplane fares
U.K.		✓ Food ✓ Optical disk ✓ PC peripherals ✓ Charges for package tours ✓ Apparel, etc.
U.S.		✓ Airplane fares ✓ Motor fuel
Australia	Not open to the public	✓ Clothing ✓ Footwear

Note: As of 2019. Based on Ministry of Internal Affairs and Communications (2019, 2021b).
Sources: Ministry of Internal Affairs and Communications (2019, 2021b).

As an example of a two-part tariff, a pricing system that combines a fixed fee (monthly) and a metered fee (charged per content) has recently been applied in the pricing of video streaming services. In this pricing system, although the subscriber can watch many videos by paying a fixed monthly fee, some content, including the latest movies, is not available without paying an additional fee.[66] Some of today's popular subscription services have not only a single two-part pricing structure, but also several two-part tariffs, depending on consumer demand patterns. A possible price survey method would be to set up representative consumers for each price plan, conduct a model price to calculate the price index based on each consumer's typical consumption pattern, and weight the obtained price index according to the percentage of expenditure. However, this method is not easy to design.

With the growing presence of e-commerce, another issue is how to incorporate a points system. At first glance, points systems are similar to rebates and frequent flyer programs, but there are unique issues. For example, since frequent flyer miles are issued for each flight and there is a one-to-one correspondence between miles and service, this mileage system can be considered a de facto discount. On the other hand, points do not necessarily correspond one-to-one with services, such as points awarded in conjunction with simultaneous purchases or non-linear point awards. Thus, even if highly granular transaction data per user, including points earned by consumers, were available, it would not be clear how to link points to the price index of an item.

There is a growing need to measure the economic value of personal data provided by consumers when using digital services. Some free services are offered by capitalizing on the value of the data (and the additional value it generates) as well as advertising fees. In order to summarize them, it is easy to understand if we classify the interaction between the service provider and the consumer as interactive or not when the user consumes the free service.

66 Theoretically, a two-part tariff is known to be complex. The optimal two-part tariff for a monopolist in a market consisting only of representative consumers is to match the pay-as-you-go charge to the marginal cost and the fixed charge to the consumer surplus, in which case the monopolist can collect all of the total surplus (Cabral, 2017). The profit maximization problem for firms under the realistic assumption that a monopolist cannot identify consumer preferences becomes more complex, and offering the same two-part tariff to all consumers with different preferences is not an optimal solution. For details, see Tirole (1988) and Ishibashi (2021) for discrete consumer preferences, and Maskin and Riley (1984) and Tirole (1988) for continuous consumer preferences. See also Wilson (1993) and Shy (2008) for more details of complex pricing models, including two-part tariffs.

One digital service for which data is one-way is broadcast (over-the-air) television services. This service is provided free of charge in the consumer market for TV broadcasting services, and the cost is charged back in the form of advertising in the business market. In return, a service is provided to the firm to deliver the opportunity for the commercial distribution on TV. In other words, a set of services is composed of a combination of transactions in two markets, one for consumers and the other for businesses.[67] In this case, as we have sufficient coverage of advertising expenditures, e.g., "television advertising (spot advertising)" in the services producer price index, there is no problem in terms of coverage in the price statistics.[68]

For example, some free services require consumers to send browsing records, etc., to service providers (e.g., Internet search services and map applications).[69] The source of revenue could include (1) cases where the advertising market is used,[70] (2) cases where firms are charged when they incorporate map data, etc., into their own paid services,[71] and (3) cases where free services are provided to attract consumers to their own service ecosystem, and the provider of the free services bears the cost of those services. In either

67 The TV distribution services illustrated here, as well as the search and payment services described below, are classified as two-sided markets. A two-sided market is a form in which trades in two markets with the network effect are connected through a specific platform. Because of network effects, there is an externality whereby the consumption of a good or service by one consumer or firm affects other consumers or firms. In a two-sided market, the network effect in one market further affects another market. For example, in TV distribution services, it is more effective for firms to advertise on channels with a large number of viewers, and in payment services, it is more efficient for merchants to install terminals that can settle payment for payment brands with a large number of users.

68 The discussion here is in terms of requirements for the deflator. In contrast, there is a separate discussion on the output side about free services not being accounted for in the current SNA (Nakamura, 2005; Nakamura and Soloveichik, 2015). In the current SNA, the cost of production (advertising costs), which is the source of the free service, is recorded as an intermediate input, thus offsetting values of the mass media side and the advertising client side in the calculation of value added in the production approach. Also, in the expenditure approach, first-order value-added is not counted because the service price is free (and a significant consumer surplus is generated in this case). In other words, only the secondary spillover effects, i.e., the sales promotion effect and the associated increase in production, would be accounted for in the SNA. These studies suggest that this calculation method should be changed to measure the added value of free services.

69 Data transmission from consumers can be mandatory or permission-based (e.g., cookie permissions).

70 In this case, search history and other information is analyzed to identify consumer preferences, which enables advertisements to be displayed only to those segments of the population that the advertiser wishes to appeal to. As mentioned in footnote 67, a network effect is realized in that the more users of search engines, etc., the more desirable it is for the firm's market.

71 Some map applications are free for consumers to use, but charge businesses when they use the functions as part of their service (e.g., when a restaurant review site/application displays a map with location information, the review site/application operator is charged). For consumers, it is easy to know the location of stores using the map application they use daily, and for businesses, it is beneficial to effectively match consumers and stores by displaying such map information.

case, the user data provided by the consumer to the firm can become a tangible or intangible asset that improves the profitability of the firm. Therefore, on the consumer side market, such services may appear to be free, but in reality this can be viewed as a trade of free services for data provided by the consumer. Theoretical studies closely related to these ideas are beginning to appear (Farboodi and Veldkamp, 2022), and if we adopt this perspective, we could, in principle, measure the economic value of the data, consider that value as a payment for the use of the service, and incorporate it into the price index.[72]

Data from payment apps and other data used by consumers can be useful in capturing the complexities of pricing faced by consumers. It could also capture the quantity and quality of goods and services consumed by households in real time, which could contribute to improving the accuracy of the CPI. In other words, by comparing price indexes using such new data sets and those following the conventional method, it is possible to evaluate not only measurement errors caused by complex pricing behavior, but also a wide range of measurement errors associated with digitalization that cannot be captured only by price surveys in actual stores. However, when introducing a new method, it is necessary to pay careful attention to the continuity of statistics and the lack of technical accumulation for the new method.

72 Although these studies are on the deflator side, the SNA revision work on standards, which corresponds to the output side, also aims at covering the free digital economy. As described in footnote 68, it has been pointed out that the current international standard, the "2008 SNA," does not sufficiently capture free digital services. The "2025 SNA" (tentative name) is considering accounting for data acquired and held by firms, including consumer purchase histories, as intangible assets like software and R&D investments, and measuring the value of free digital services such as Internet searches and social networking services. The general idea is that data storage and database construction can be regarded as non-marketable items, so the method of calculating their value by accumulating their costs (labor costs) is discussed (Economic and Social Research Institute in the Cabinet Office, 2023; Statistics Canada, 2019; de Bondt and Mushkudiani, 2021). However, the following issues remain: how to handle data with a usage period of less than one year (outside the definition of gross fixed capital formation); whether to include data analysis in the aggregate; how to define the scope of labor cost aggregation when data processing is performed by non-professionals; how to calculate depreciation rates; and how to capture international transactions with overseas entities (Economic and Social Research Institute in the Cabinet Office, 2023).

References

Papers in Japanese Only

Abe, Naohito, *Bukka Shisu Gairon: Shisu/Shukei Riron heno Shotai* (Introduction to Price Indexes: Invitation to Index and Aggregation Theory), Nippon Hyoron Sha, 2023 (in Japanese).

Bank of Japan, "Hi-dentouteki Kinyu Seisaku to Infure Yoso (Unconventional Monetary Policy and Inflation Expectation)," the Second Workshop on the "Review of Monetary Policy from a Broad Perspective," Monetary Affairs Department, Bank of Japan, 2024
(available at https://www.boj.or.jp/mopo/outline/bpreview/data/bpr240405a2.pdf,
Last access date, same as below: May 21, 2024, in Japanese).

Economic and Social Research Institute in the Cabinet Office, "2025SNA (Kasho) nimuketa Digital Economy no Keisoku nikansuru Chosa Kenkyu: Data no Shihon tositeno Kiroku Hoho nitsuite (Research on Measurement of Digital Economy for 2025SNA (tentative name): How to Record Data as Capital)," Reports of Study Groups, 88, 2023
(available at https://www.esri.cao.go.jp/jp/esri/prj/hou/hou088/hou88a.pdf,
January 24, 2024, in Japanese).

Ishibashi, Koji, *Industrial Organization: A Synthesis of Theory and Empirics*, Keio University Press, 2021 (in Japanese).

Ministry of Internal Affairs and Communications, "Kouri Bukka Tokei Chosa Kakaku Chosa Chiku no Settei Hoho nitsuite (Methodology for Setting Price Survey Districts in Retail Price Survey)," Statistics Bureau, Ministry of Internal Affairs and Communications, 2003
(available at https://www.stat.go.jp/data/kouri/doukou/pdf/minaoshi.pdf,
January 24, 2024, in Japanese).

———, "Shohisha Bukka Shisu (CPI) heno Web Scraping no Katsuyo nitsuite (Use of Web Scraping for the Consumer Price Index [CPI])," Price Statistics Office, Statistics Bureau, Ministry of Internal Affairs and Communications, 2019
(available at https://www.soumu.go.jp/main_content/000654895.pdf,
January 24, 2024, in Japanese).

———, “Shohisha Bukka Shisu niokeru Minei Yachin no Keinen Henka no Chosei Hoho (How to Adjust for Ageing of Private Rents in the Consumer Price Index),” Price Statistics Office, Statistics Bureau, Ministry of Internal Affairs and Communications, 2021a
(available at https://www.stat.go.jp/data/cpi/pdf/kenkyu2.pdf,
January 24, 2024, in Japanese).

———, “Web Scraping niyoru Hinmokubetsu Kakaku Shisu no Sanshutsu (Calculation of Price Index by Item via Web Scraping),” Statistics Bureau, Ministry of Internal Affairs and Communications, 2021b
(available at https://www.stat.go.jp/data/cpi/2020/kaisetsu/pdf/fu3.pdf,
January 24, 2024, in Japanese).

Nishimura, Kiyohiko G., and Masahiro Higo, “Post-corona Jidai no Kouteki Tokei (2): Bukka/Chingin Tokei no Seido Mondai (Official Statistics in the Post Covid-19 Era (2): Accuracy Problems of Price and Wage Statistics),” CREPE Discussion Paper No.136, Center for Research and Education for Policy Evaluation, University of Tokyo, 2022 (in Japanese).

———, Nariyasu Yamasawa, and ———, *Tokei ―Kiki to Kaikaku: System Rekka karano Fukkatsu* (Statistics Crisis and Reform: Recovery from Systemic Deterioration), Nikkei Publishing, 2020 (in Japanese).

Nishizaki, Toshimi, and Susumu Kuwahara, “Research on a Method for Estimating Prices Reflecting Changes in the Quality of Medical Care,” *Keizaibunseki* (*Economic Analysis*), 207, Economic and Social Research Institute in the Cabinet Office, 2023, pp.220-249 (in Japanese).

Onaka, Yuichi, “Aged Depreciation Adjustment of Rent in the Consumer Price Index,” Research memoir of the statistics, 79, Ministry of Internal Affairs and Communications, 2022, pp.93-110 (in Japanese).

Ota, Makoto, *Hinshitsu to Kakaku: Atarashii Shouhisha no Riron to Keisoku* (Quality and Price: New Consumer Theory and Measurement), Sobun Sha, 1980 (in Japanese).

Shiratsuka, Shigenori, *Bukka no Keizai Bunseki* (Economic Analysis of Prices), University of Tokyo Press, 1998 (in Japanese).

———, “Shohisha Bukka Shisu no Seido Kojo nimukete: Choki ni Watari Tsumi Nokosareteiru Kadai no Saikento (Toward Improving the Accuracy of the Consumer Price Index: Reexaminations of Longstanding Issues),” in Watanabe, Tsutomu, and Chihiro Shimizu, eds. *Price Dynamics in Goods and Asset Markets*, University of Tokyo Press, 2023 (in Japanese).

———, and Sachiko Kuroda, “Video Camera Kakaku no Hedonic Bunseki (Hedonic Analysis of Video Camera Prices),” *Kinyu Kenkyu*, 14(4), Institute for Monetary and Economic Studies, Bank of Japan, 1995, pp.43-62 (in Japanese).

Suga, Mikio, *Bukka Shisu no Sokutei Ron: Micro Data niyoru Keiryo Keizaigakuteki Sekkin* (The Measurement Theory of Price Index: An Econometric Approach Using Micro Data), Nippon Hyoron Sha, 2005 (in Japanese).

Umeda, Masanobu, “Nihon no Shohisha Bukka Shisu no Sho Tokusei to Kinyu Seisaku Unei (Various Characteristics of Japan's CPI and Monetary Policy Management),” in Yoshikawa, Hiroshi, ed. *Defure Keizai to Kinyu Seisaku* (Deflationary Economy and Monetary Policy), Keio University Press, 2009 (in Japanese).

Yoshida, Jiro, “Genkai Yachin Shisu no Suikei: Shouhisha Bukka Shisu no Kaizen nimukete (Marginal Rent Index: Toward an Improved Consumer Price Index),” in Watanabe, Tsutomu, and Chihiro Shimizu, eds. *Price Dynamics in Goods and Asset Markets*, University of Tokyo Press, 2023 (in Japanese).

Yoshizoe, Yasuto, “Some Challenging Issues in Economic Indicators,” *Nihon Tokei Gakkai-shi*, 46(2), 2017, pp.173-191 (in Japanese).

Papers in English, etc.

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

Aizcorbe, Ana M., and Patrick C. Jackman, "The Commodity Substitution Effect in CPI Data, 1982-91," *Monthly Labor Review*, 116(12), Bureau of Labor Statistics, 1993, pp.25-33.

Alchian, Armen A., and Benjamin Klein, "On a Correct Measure of Inflation," *Journal of Money, Credit and Banking*, 5(1), 1973, pp.173-191.

Ambrose, Brent W., N. Edward Coulson, and Jiro Yoshida, "The Repeat Rent Index," *Review of Economics and Statistics*, 97(5), 2015, pp.939-950.

———, ———, and ———, "Reassessing Taylor Rules Using Improved Housing Rent Data," *Journal of Macroeconomics*, 56, 2018, pp.243-257.

———,———, and ———, "Housing Rents and Inflation Rates," *Journal of Money, Credit and Banking*, 55(4), 2023, pp.975-992.

Aoki, Shuhei, and Minoru Kitahara, "Measuring a Dynamic Price Index Using Consumption Data," *Journal of Money, Credit and Banking*, 42(5), 2010, pp.959-964.

Ariga, Kenn, and Kenji Matsui, "Mismeasurement of the CPI," in Magnus Blomström, Jennifer Corbett, Fumio Hayashi, and Anil Kashyap, eds. *Structural Impediments to Growth in Japan*, University of Chicago Press, 2003.

Australian Bureau of Statistics, "CPI International Comparisons," Australian Bureau of Statistics, 2023 (available at https://www.abs.gov.au/articles/cpi-international-comparisons, January 24, 2024).

Bean, Charles, "Independent Review of UK Economic Statistics: Final Report," Independent Report, Government of the United Kingdom, 2016 (available at https://assets.publishing.service.gov.uk/media/5a7f603440f0b62305b86c45/2904936_Bean_Review_Web_Accessible.pdf, January 24, 2024).

Bils, Mark, and Peter J. Klenow, "Quantifying Quality Growth," *American Economic Review*, 91(4), 2001, pp.1006-1030.

Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale Jorgenson,"Toward a More Accurate Measure of the Cost of Living: Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index," Advisory Commission to Study the Consumer Price Index, 1996.

Bradley, Ralph, "Feasible Methods to Estimate Disease Based Price Indexes," *Journal of Health Economics*, 32(3), 2013, pp.504-514.

Broda, Christian, and David E. Weinstein, "Defining Price Stability in Japan: A View from America," *Monetary and Economic Studies*, 25(S-1), 2007, pp.169-206.

Brown, Craig, and Anya Stockburger, "Item Replacement and Quality Change in Apparel Price Indexes," *Monthly Labor Review*, 129(12), Bureau of Labor Statistics, 2006, pp.35-45.

Bundesbank, "Changes in the Official Consumer Price Statistics and Their Implications for the 'Measurement Bias' in the Inflation Rate," *Monthly Report*, 54(8), Deutsche Bundesbank, 2002, pp.38-39.

Byrne, David, Carol Corrado, and Daniel Sichel, "The Rise of Cloud Computing: Minding Your Ps, Qs and Ks," in Carol Corrado, Jonathan Haskel, Javier Miranda, and Daniel Sichel, eds. *Measuring and Accounting for Innovation in the Twenty-First Century*, University of Chicago Press, 2021.

Cabral, Luis M. B., *Introduction to Industrial Organization, second edition*, MIT Press, 2017.

Cage, Robert, John Greenlees, and Patrick Jackman, "Introducing the Chained Consumer Price Index," paper presented at the Seventh Meeting of the International Working Group on Price Indices in Paris, 2003.

Carli, Gian R., *Del Valore e della Proporzione dei Metallic Monetati con i Generi in Italia Prima delle Scoperte dell'Indie col Confronto del Valore e della Proporzione de'Tempi Nostri, Lucca*, Reprinted in a Series Dedicated to the Italian Economists in 1804, 1764.

Cecchetti, Stephen, "Housing in Inflation Measurement," VOXEU Column, Center for Economic Policy Research, 2007 (available at https://cepr.org/voxeu/columns/housing-inflation-measurement, January 24, 2024).

Collin, Marianne, "International Methodological Standards for CPI and National Practices," *IFC Bulletin*, 24, Irving Fisher Committee on Central Bank Statistics, 2006, pp.19-32.

Covas, Francisco, and João S. Silva, "Outlet Substitution Bias," Economic Bulletin, Banco de Portugal, 1999.

CPI Commission, "The CPI Commission: Discussion," *American Economic Review*, 87(2), 1997, pp.94-98.

Crone, Theodore M., Leonard I. Nakamura, and Richard Voith, "Rents Have Been Rising, Not Falling, in the Postwar Period," *Review of Economics and Statistics*, 92(3), 2010, pp.628-642.

Cutler, David M., Mark McClellan, Joseph P. Newhouse, and Dahlia Remler, "Are Medical Prices Declining ?" NBER Working Paper No.5750, National Bureau of Economic Research, 1996.

———, ———, ———, and ———, "Are Medical Prices Declining? Evidence from Heart Attack Treatments," *Quarterly Journal of Economics*, 113(4), 1998, pp.991-1024.

Darby, Michael R., "Statement," in *Consumer Price Index: Hearings before the Committee on Finance,United States Senate*, U.S. Government Printing Office, 1995, pp.173-176.

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

de Bondt, Hugo, and Nino Mushkudiani, "Estimating the Value of Data in the Netherlands," presented at the IARIW-ESCoE Conference 'Measuring Intangible Assets and Their Contribution to Growth' in London, 2021.

Diewert, W. Erwin, "Exact and Superlative Index Numbers," *Journal of Econometrics*, 4(2), 1976, pp.115-145.

———, "Superlative Index Numbers and Consistency in Aggregation," *Econometrica*, 46(4), 1978, pp.883-900.

———, "Aggregation Problems in the Measurement of Capital," in Dan Usher eds. *The Measurement of Capital*, University of Chicago Press, 1980, pp.433-538.

———, "Prepared Statement," in *Consumer Price Index: Hearings before the Committee on Finance,United States Senate*, U.S. Government Printing Office, 1995, pp.115-118.

———, "Index Number Issues in the Consumer Price Index," *Journal of Economic Perspectives*, 12(1), 1998, pp.47-58.

———, and Alice O. Nakamura, "Accounting for Housing in a CPI," in W. Erwin Diewert, Bert M. Balk, Dennis Fixler, Kevin J. Fox, and Alice O. Nakamura, eds. *Price and Productivity Measurement: Volume 1- Housing*, Trafford Press, 2009.

Drobisch, Moritz Wilhelm, "Über Mittelgrossen und die Anwendbarkeit derselben auf die Berechnung des Steigens und Sinkens des Goldwerthes," Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Bd.XXIII, 1871, pp.25-48.

Dutot, Nicolas, *Reflexions Politiques sur les Finances et le Commerce,* A La Haye: Vaillant and Nicolas Prévost, 1738.

Eiglsperger, Martin, Rodolfo Arioli, Bernhard Goldhammer, Eduardo Gonçalves, and Omiros Kouvavas, "Owner-Occupied Housing and Inflation Measurement," published as part of the ECB Economic Bulletin Issue 1/2022, European Central Bank, 2022.

European Central Bank (ECB), "Inflation Measurement and Its Assessment in the ECB's Monetary Policy Strategy Review," Occasional Paper Series No.265, European Central Bank, 2021.

Farboodi, Maryam, and Laura Veldkamp, "A Model of the Data Economy," NBER Working Paper No.28427, National Bureau of Economic Research, 2022.

Fisher, Irving, *The Making of Index Numbers: A Study of Their Varieties, Tests, and Reliability,* Houghton Mifflin Company, 1922.

Funke, Helmut, Günther Hacker, and Joachim Voeller, "Fisher's Circular Test Reconsidered," *Schweizerische Zeitschrift für Volkswirtshaft und Statistik*, 115(IV), 1979, pp.677-688.

General Accounting Office (GAO), "Consumer Price Index: Update of Boskin Commission's Estimate of Bias," GAO/GGD-00-50, U.S. General Accounting Office, 2000.

Goolsbee, Austan D., and Peter J. Klenow, "Internet Rising, Prices Falling: Measuring Inflation in a World of E-Commerce," *AEA Papers and Proceedings*, 108, 2018, pp.488-492.

Gordon, Robert J., *The Measurement of Durable Goods Prices,* University of Chicago Press, 1990.

———, "Prepared Statement," in *Consumer Price Index: Hearings before the Committee on Finance,United States Senate*, U.S. Government Printing Office, 1995, pp.122-126.

———, "The Boskin Commission: A Retrospective One Decade Later," NBER Working Paper No.12311, National Bureau of Economic Research, 2006.

———, and Todd vanGoethem, "Downward Bias in the Most Important CPI Component: The Case of Rental Shelter, 1914-2003," in Ernst R. Berndt and Charles R. Hulten, eds. *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, University of Chicago Press, 2007, pp.153-195.

Greenlees, John S., and Robert McClelland, "New Evidence on Outlet Substitution Effects in Consumer Price Index Data," *Review of Economics and Statistics*, 93(2), 2011, pp.632-646.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell, "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, 87(3), 1997, pp.342-362.

Griliches, Zvi, "Prepared Statement," in *Consumer Price Index: Hearings before the Committee on Finance,United States Senate*, U.S. Government Printing Office, 1995, pp.129-132.

Hausman, Jerry, "Cellular Telephone, New Products, and the CPI," *Journal of Business & Economic Statistics*, 17(2), 1999, pp.188-194.

———, and Ephraim Leibtag, "CPI Bias from Supercenters: Does the BLS Know that Wal-Mart Exists?" in W. Erwin Diewert, John S. Greenlees, and Charles R. Hulten, eds. *Price Index Concepts and Measurement*, University of Chicago Press, 2009.

Higo, Masahiro, and Shigenori Shiratsuka, "Consumer Price Measurement under the First Wave of the COVID-19 Spread in Japan: Scanner Data Evidence for Retailers in Tokyo," *Japan and the World Economy*, 65(101176), 2023.

Hoffmann, Johannes, "Problems of Inflation Measurement in Germany," Discussion Paper No.1/98, Deutsche Bundesbank, 1998.

———, "Problems of Inflation Measurement in Germany: An Update," paper presented at the Eurostat Conference on the Measurement of Inflation in Cardiff, 1999.

International Monetary Fund (IMF), International Labour Organization, Statistical Office of the European Union (Eurostat), United Nations Economic Commission for Europe, Organisation for Economic Co-operation and Development, and the World Bank, *Consumer Price Index Manual: Concepts and Methods*, 2020.

Jevons, William S., *A Serious Fall in the Value of Gold Ascertained and Its Social Effects Set Forth*, Edward Stanford, 1863.

Johnson, David S., Stephen B. Reed, and Kenneth J. Stewart, "Price Measurement in the United States: A Decade after the Boskin Report," *Monthly Labor Review*, 129(5), Bureau of Labor Statistics, 2006, pp.10-19.

Jorgenson, Dale W., "Statement," in *Consumer Price Index: Hearings before the Committee on Finance, United States Senate*, U.S. Government Printing Office, 1995, pp.36-41.

Kliesen, Kevin L., "Critiquing the Consumer Price Index," Regional Economist, Federal Reserve Bank of St. Louis, 1997 (available at https://www.stlouisfed.org/publications/regional-economist/july-1997/critiquing-the-consumer-price-index, January 24, 2024).

Lancaster, Kelvin J., "A New Approach to Consumer Theory," *Journal of Political Economy*, 74(2), 1966, pp.132-157.

Lane, Walter F., William C. Randolph, and Stephen A. Berenson, "Adjusting the CPI Shelter Index to Compensate for Effect of Depreciation," *Monthly Labor Review*, 111(10), Bureau of Labor Statistics, 1988, pp.34-37.

Laspeyres, Etienne, "Die Berechnung einer mittleren Waarenpreissteigerung," *Jahrbücher für Nationalökonomie und Statistik*, 16, 1871, pp.296-315.

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

———, John M. Roberts, and David J. Stockton, "Monetary Policy and 'The Price Level'," mimeo, Board of Governors of the Federal Reserve System, 1994.

Leifer, Hans-Albert, "Zur Berechnung von Preismeßzahlen auf der 'Elementarebene' des Harmonisierten Verbraucherpreisindex (HVPI)," [About the Calculation of "Elementary Aggregate Indices"of the HICP], *Allgemeines Statistisches Archiv*, 83, 1999, pp.338-349.

Lequiller, François, "Does the French Consumer Price Index Overstate Inflation?" Série des documents de travail de la Direction des Etudes et Synthèses Économiques No.G 9714, Insee, 1997.

Maskin, Eric, and John Riley, "Monopoly with Incomplete Information," *RAND Journal of Economics*, 15(2), 1984, pp.171-196.

Matsumoto, Brett, "Producing Quality Adjusted Hospital Price Indexes," BLS Working Paper No.543, Bureau of Labor Statistics, 2021.

McClelland, Robert, and Marshall Reinsdorf, "Small Sample Bias in Geometric Mean and Seasoned CPI Component Indexes," BLS Working Paper No.324, Bureau of Labor Statistics, 1999.

Moulton, Brent R., "Basic Components of the CPI: Estimation of Price Changes," *Monthly Labor Review*, 116(12), 1993, pp.13-24.

———, "The Measurement of Output, Prices, and Productivity: What's Changed since the Boskin Commission?" Brookings Institution, 2018.

———, and Karin E. Smedley, "A Comparison of Estimators for Elementary Aggregates of the CPI," paper presented at Western Economic Association International Conference in San Diego, 1995.

Nakamura, Leonard I., "Advertising, Intangible Assets, and Unpriced Entertainment," Working Paper No.05-11, Federal Reserve Bank of Philadelphia, 2005.

———, and Rachel H. Soloveichik, "Valuing 'Free' Media across Countries in GDP," Working Paper No.15-25, Federal Reserve Bank of Philadelphia, 2015.

Paasche, Hermann, "Über die Preisentwicklung der Letzten Jahre Nach den Hamburger Börsennotirungen," *Jahrbücher für Nationalökonomie und Statistik*, 12, 1874, pp.168-178.

Reinsdorf, Marshall, "The Effect of Outlet Price Differentials on the U.S. Consumer Price Index," in Murray Foss, Marylin Manser, and Allan Young, eds. *Price Measurements and Their Uses*, University of Chicago Press, 1993, pp.227-258.

———, and Brent R. Moulton, "The Construction of Basic Components of Cost-of-Living Indexes," in Timothy F. Bresnahan and Robert J. Gordon, eds. *The Economics of New Goods*, University of Chicago Press, 1996, pp.397-436.

———, and Paul Schreyer, "Measuring Consumer Inflation in a Digital Economy," paper presented at the Fifth IMF Statistical Forum, 2017 (available at https://www.imf.org/en/News/Seminars/Conferences/2017/05/03/5th-statistical-forum, January 24, 2024).

Reis, Ricardo, "A Dynamic Measure of Inflation," NBER Working Paper No.11746, National Bureau of Economic Research, 2005.

Rosen, Sherwin, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82(1), 1974, pp.34-55.

Rossiter, James, "Measurement Bias in the Canadian Consumer Price Index," Working Paper No.2005-39, Bank of Canada, 2005.

Sabourin, Patrick, "Measurement Bias in the Canadian Consumer Price Index: An Update," Bank of Canada Review Summer 2012, Bank of Canada, 2012, pp.1-11.

Sato, Kazuo, "Ideal Index Numbers that Almost Satisfy the Factor Reversal Test," *Review of Economics and Statistics*, 56(4), 1974, pp.549-552.

———, "The Ideal Log-change Index Number," *Review of Economics and Statistics*, 58(2), 1976, pp.223-228.

Shapiro, Matthew D., and David W. Wilcox, "Mismeasurement in the Consumer Price Index: An Evaluation," *NBER Macroeconomics Annual*, 11, 1996, pp.93-154.

Shiratsuka, Shigenori, "Effects of Quality Changes on the Price Index: A Hedonic Approach to the Estimation of a Quality Adjusted Price Index for Personal Computers in Japan," *Monetary and Economic Studies*, 13(1), 1995a, pp.17-52.

———, "Automobile Prices and Quality Changes: A Hedonic Price Analysis of Japanese Automobile Market," *Monetary and Economic Studies*, 13(2), 1995b, pp.1-44.

———, "Measurement Errors in the Japanese Consumer Price Index," *Monetary and Economic Studies*, 17(3), 1999, pp.69-102.

———, "Measurement Errors in the Japanese CPI," *IFC Bulletin*, 24, Irving Fisher Committee on Central Bank Statistics, 2006, pp.36-43.

———, "Comments on 'Defining Price Stability in Japan: A View from America' by Christian Broda and David E. Weinstein," Summary Report: ESRI International Conference "Japan's Economy and Macroeconomic Policies from 1980 to 2007," 2007 (available at https://warp.da.ndl.go.jp/info:ndljp/pid/11539153/www.esri.go.jp/en/workshop/070625/070625gaiyou-e.html, January 24, 2024).

Shy, Oz, *How to Price: A Guide to Pricing Techniques and Yield Management*, Cambridge University Press, 2008.

Statistics Canada, "The Value of Data in Canada: Experimental Estimates," Latest Developments in the Canadian Economic Accounts, Statistics Canada, 2019.

Takahashi, Wataru, Kenichiro Watanabe, and Hiroshi Fujiki, "The Fourteenth International Conference: Growth, Integration and Monetary Policy in East Asia," *Monetary and Economic Studies*, 25(S-1), 2007, pp.1-8.

Takahashi, Yuta, and Naoki Takayama, "Global Technology Stagnation," mimeo, 2023.

Tirole, Jean, *The Theory of Industrial Organization*, MIT Press, 1988.

Ueda, Kozo, Kota Watanabe, and Tsutomu Watanabe, "Price Setting in Online and Offline Markets: Evidence from Korea," paper presented at Seventeenth Ottawa Group Meeting in Rome, 2022.

Vartia, Yrjö O., "Ideal Log-change Index Numbers," *Scandinavian Journal of Statistics*, 3(3), 1976, pp.121-126.

Wilson, Robert B., *Nonlinear Pricing*, Oxford University Press, 1993.

Wynne, Mark A., "An Estimate of the Measurement Bias in the HICP," Working Paper No.0509, Federal Reserve Bank of Dallas, 2005.