NBER WORKING PAPER SERIES

MISMEASUREMENT OF THE CPI

Kenn Ariga
Kenji Matsui

Working Paper 9436
http://www.nber.org/papers/w9436

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2003

This paper was prepared for the volume on Structural Impediments to Growth that is being edited by Magnus Blomstrom, Jennifer Corbett, Fumio Hayashi and Anil Kashyap for the NBER and will be published by the University of Chicago Press. The authors benefited greatly from valuable comments and suggestions by Albert Ando, Jenny Corbett, David Flath, Fumio Hayashi, Yasushi Iwamoto, Anil Kashyap, Robert Lipsey, Shigenori Shiratsuka, and other participants in the NBER Japan Project. We also wish to thank Ignazio Visco and Enrico Giovanni at OECD for helping locate data on statistical programs of OECD countries. Special thanks are due to Mr Masato Aida and others at Japan's Statistics Bureau for their replies to our queries regarding the price survey procedures used in the CPI and for data on the educational background of statistical staff. Finally, Larry Meisner's superb professional editing has made the presentation in the final version far clearer. The authors are of course responsible for any errors in the paper. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Mismeasurement of the CPI
Kenn Ariga and Kenji Matsui
NBER Working Paper No. 9436
January 2003
JEL No. F4, F1

ABSTRACT

In this paper, we investigate several key problems in Japanese economic statistics. We use CPI mismeasurements and biases as an example to explore the roots of the problems and also to offer guidelines for improvements. We emphasize 3 major shortcomings shared by many official statistics in Japan: (1) long delays in adjustments, (2) lack of proper coordination, and 3) insufficient information disclosure.

In the analysis of CPI bias, we limit our focus to potential biases due to aggregation, survey methodology and sample selection procedures. We estimate that, in recent years, the commodity CPI inflation rate is biased upward by at least 0.5% per year, even if we assume away the potential bias associated with the quality adjustment, delay in incorporating changes in consumption basket, and other important unresolved problems.

Kenn Ariga
Kyoto Institute of Economic Research
Kyoto University
Yoshida-Honmachi, Sakyo-ku
Kyoto 606 Japan
ariga@kier.kyoto-u.ac.jp

Kenji Matsui
Faculty of Business Administration
Yokohama National University
79-4 Tokiwa-dai Hodogaya-ku
Yokohama, Kanagawa 2408501
Japan
mkenzi@white.plala.or.jp

As the Japanese economy continues to experience negative or near-zero growth under weak demand, many economists and policy makers are increasingly concerned over the accuracy of many key economic statistics. In particular, the accuracy of the CPI (consumer price index) has become a central issue.

The annual CPI registered declines in the three years through 2001. In early 2002, the data indicate the possibility deflation might be somewhat accelerating. While the economy seems to be floating at the edge of a deflationary spiral, many suspect and are worried that prices are falling faster than CPI statistics suggest. Supporting these concerns are such things as Seiyu, a large supermarket chain, publishing an index showing how its own prices had fallen much faster than the official CPI.

If CPI data contain significant measurement errors, such that the downward trend is not measured with accuracy, the cost of such bias can be substantial. Consider, for example, potential ramifications on the heated debate over monetary policy, especially inflation targeting. The very idea of inflation targeting hinges critically on timely and accurate measurement of the inflation rate. Because retail price data collected by Sōmusho (Ministry of Public Management, Home Affairs, Posts and Telecommunications) for the CPI are also used for the national income statistics, mis-measurements in the CPI can lead to serious errors in GDP statistics as well.

In general, the potential cost of mis- or non-measurements in official statistics can be substantial and are not limited to affecting policy making. Many economists in the financial sector and consulting firms have voiced concern over the noise and inconsistency in the quarterly GDP estimates. The discrepancy between preliminary and final GDP figures is suspected to originate in the inconsistency in several dimensions in the methodologies employed in the two estimates.[1]

---

[1] For example, in the preliminary GDP figures, private fixed investment growth for year 2000 initially was reported as 4.6%; this became 9.3% in the final figure. For a brief review of the quality-of-statistics issue, and a response by the Economic and Social Research Institute of Cabinet Office, visit [www5.cao.go.jp/2000/g/0602g-gdpcoments.html].

Given the critical role of key economic indicators such as GDP and CPI, it is not surprising that large and frequent swings in official statistics can create visible commotions in financial markets and other sectors of the economy. The potential costs due to problems in the official statistics are widespread and far reaching. For example, a key part of the "structural reform" advanced by the Koizumi government is job creation in services and information technologies (IT). However, there are no official statistics to guide such policy, as none of the published data report job creation by start-ups or job destruction from closing of establishments.

To be fair, there are many good, even wonderful, things to say about economic statistics in Japan. There is extensive and comprehensive coverage on a wide spectrum of topics, especially those collected on an establishment basis. Some are quite exotic and probably not available anywhere else in the world. Many are collected by non-government institutions. Moreover, data are comprehensive, geographically and otherwise. Although the country consists of many small islands, most government statistics cover virtually the entire population.

There are problems, of course, some rather serious in nature and quantitatively important. In this paper, we point out several underlying factors responsible for the problems in official statistics in Japan. In doing this, however, the focus is on the consumer price index (CPI). Most of the problems raised by the Boskin Commission for the US CPI are found in Japan's CPI. In many areas, the potential ramifications seem even more important in Japan.

CPI is an important and popular statistic and is used for many different purposes. The CPI inflation rate is one of the key indicators for cyclical fluctuations of the economy. CPI also is used as the benchmark in many wage setting negotiations and public pensions are linked to it. (Although for political reasons the pensions have not been adjusted downward to reflect the sizable decline in the CPI.) Recent macroeconomic developments in Japan also add to the significance of studying potential mismeasurements in CPI in that the stagnant economy has been experiencing zero or negative inflation rates for prolonged periods, an experience that is rather unique and which might shed new light on issues of measurement biases in CPI.

The choice of the CPI is partly because of the authors' background: in past research we have used disaggregated price data as well as price indices such as CPI or WPI and we are, therefore, concerned about their accuracy. More important is the depth of the analysis that can be achieved. Thus, although GDP is by far the most popular and important statistic, it is a secondary one based on a large variety of primary statistics. That means the potential sources of biases and other problems are simply too great to be thoroughly analyzed in a single paper.[2]

Moreover, the CPI shares with other major official statistics the underlying causes which lie beneath the problems in the Japanese official statistics system. We hope this investigation of the CPI helps elucidate the nature of the problems commonly found in many important official economic statistics of Japan.

The paper is organized as follows. First we offer a bird's eye view of official statistics in Japan and point out several important deficiencies, after which we review key issues in the CPI. With this background, we investigate potential problems in several major aspects of the CPI. These include data collection procedures (including how discounted prices are handled), services, quality change and new products, and aggregation issues (substitution across time, brands, and stores). We then look at a discrepancy between CPI and WPI that probably relates to differences in how quality adjustments are made and some hitherto neglected aspects of the measurement problem, relating to shopping and storage behaviors. From this analysis we offer a tentative assessment of the magnitude of the CPI inflation rate bias and draw some suggestions for improving the statistics in general and the CPI in particular.

A cautionary note on the distinction between potential measurement errors in general and bias in the inflation rate: measurement errors contaminate the CPI, but they do not imply systematic bias in the measured inflation rate, or changes in COLI. For example, consider medical and health-care services. Although we believe there are serious measurement errors and under-representation problems, it is unclear if and in which direction they affect the

---

[2] For the revised System of National Account (SNA) in Japan, see: http://www.esri.cao.go.jp/en/sna/020612/outline.pdf. Ando(2002) explains in great details the problems he encountered in SNA data as he investigates the cause of the long stagnation of the Japanese Economy. For those not familiar with Japanese economic statistics, Matsuoka and Rose (1994) provides a gateway into major economic statistics in Japan.

measured inflation rate. Indeed, several indices shown in Iwamoto (2000) indicate higher, and others, lower, inflation in medical expenditure than does the CPI.

1 Overview of Official Statistics

Japanese official statistics fall into three broad groups based on how they are created. Primary statistics collected for specific purposes (*chōsa-tōkei*), primary statistics collected as part of the regular tasks of governmental offices (*gyōmu-tōkei*), and processed statistics derived from primary data. Primary data on exports and imports (Custom Clearance Statistics) compiled by the Ministry of Finance (MOF) is an important example of the second group. The National Accounts are by far the most well known of the last group.

A more important distinction among *chōsa-tōkei* is based on legal status. The core of official statistics are called "designated statistics" (*shitei-tōkei*). There also are "approved statistics," so named because they are approved by the Minister of Sōmusho (Minister of Public Management, Posts and Telecommunications).

Designated and approved statistics have special status in the law. Specifically, the law stipulates clearly that government bodies collecting these statistics are endowed with authority to request and enforce proper cooperation from the public chosen to be surveyed. At the same time, the law sets rather rigid restrictions on the use and dissemination of information so obtained. This allows the data collecting agency to conduct surveys and census in a way that private bodies without such authorization cannot hope to accomplish. In short, compared to other official statistics, these two types of statistics are given priority in data collection and a more stringent set of rules governs their use and dissemination.

Table 1 lists the number of designated and approved statistics by the ministry responsible for collecting them, with the ministry's staff and budget for statistics.

1.1 Staffing and Collecting

Officially, the Statistics Bureau of Sōmusho is responsible for coordinating the activities of the statistics sections of all ministries. It is apparent, however, that the system is highly decentralized and each ministry seems to act on its own in creating, collecting, abandoning, and publishing data. Which ministry is responsible for a series often is a historical accident, but ministries seem unwilling to reshuffle assignments. For example, Sōmusho conducts the Survey of Research and Development, the National Tax Agency collects data on salaries in the private sector, and the Bank of Japan (technically not even a part of the government) compiles the wholesale price index (WPI) and corporate service price index (CSPI).

Table 2 displays data for the US federal government comparable to Table 1. Its budget in 2002 was roughly 10 times that of Japan in absolute terms, and over 3 times in share terms. Although total staff is similar, this is only because of the large number employed at Japan's Ministry of Agriculture, Forestry and Fishery.

Composition of staff in Japan is problematic. As far as is known, only a very few workers actually have advanced degrees in statistics, and virtually no one does in economics. Based upon information from Sōmusho, Statistics Bureau, perhaps 10 (out of 384 full time staff) have an MS in statistics, and no one had an MA or PhD in economics[3]. In contrast, the US federal government employs more than 2,000 professional statisticians on a full time permanent basis (It is not clear, however, how many of them have advanced degrees in statistics or economics. In any case, we are certain that US government professional staff with advanced degrees far outnumbers that in the Japanese counterpart).

2 General Data Problems

Japanese statistics have several broad problems in addition to the absence of statistical professionals among the staff mentioned above. These include long lead times, coordination among agencies, appropriateness of the data collected, and access to raw data and information on how data are processed.

[3] We are grateful to Mr Masato Aida at Sōmusho for this information.

2.1 Long Delays in Adjustments

Titles of the designated statistics indicate that their coverage is far from being well balanced. Although each series differs in scope and size, Table 1 is at least suggestive of the imbalance between the coverage of official statistics and the relative importance of subjects covered. This reflects slowness in changing the data collected to reflect changes in the economy. The imbalance is particularly noticeable in agriculture and fishery. In 1999, the Ministry of Agriculture, Forestry and Fisheries spent 29% of the total budget and employed 68% of staff devoted to statistics collection and compilation. But all primary industries combined provide less than 2% of GDP.

Some other examples are: domestic production and usage of coal is a designated series even though only 1.9% of total coal consumption is produced domestically and only 12% of total energy consumptions is coal. There are three designated statistics on shipping and sailors, although Japanese commercial ships long ago replaced Japanese crew with foreigners. Even though the industry was all but extinct years earlier, production of silk and silk worms was a designated statistic until the end of fiscal 2002.

On the other hand, surprisingly few resources are allocated for data on tertiary industries, especially services. There is only one designated series that covers the service industry on an annual basis, offering basic data on production, employment, firm size, etc. Even this statistic rotates among subsectors on a three-year cycle so that the data for each subsector is available only every third year. There is only one other designated series that covers the service industry but this survey is conducted every five years and it covers only those not covered in the first survey.

As we see more closely later, the weights attached to items in the CPI are based on the FIES (Family Income and Expenditure Survey), and it is fixed for a five-year period, even though the FIES is conducted monthly. Japan is not unique in this, other countries also have

similar delays in adjusting coverage and weights. In the US CPI, 1982-84 weights were used until 1996, finally being replaced by 1993-95 weights.

Especially for GDP statistics, long lead times are a problem. Preliminary figures are not announced until three months after the end of a quarter. These are revised three months later. The final figure is made available in December of the next year. Moreover, the inconsistency between quarterly estimates and the final figures reflects underlying differences in the estimation procedure. The inconsistency and long lead times in GDP statistics have been known for quite some time, but there seems little hope that any fundamental measures will be taken to rectify the situation. In the United States preliminary quarterly GDP data are announced in 8 weeks and the final figure is available in about 13 weeks. In other words, by the time the preliminary Japanese figures are announced, the final US figure is announced. The release of the latest CPI figures is far more timely. The most recent month's figure is released on the Friday of the last week of each month, whereas CPI in the second ten days of the current month for metropolitan Tokyo area is released on the same day.

2.2 Lack of Proper Coordination

There is a lack of proper coordination among different bodies of government and coordination with non-governmental institutions is uncommon. As a result, different bodies collect similar, if not duplicate, sets of data. At the same time, in many important areas there is a lack of proper official statistics, due, mainly, to the fact that the area falls under more than one ministry's responsibility. This is especially true in the areas of information and communication: subsets of these are covered rather independently by sections of Sōmusho and the Ministry of Economy, Trade and Industry (formerly, MITI).

Inadequate coordination creates difficulties in combining sets of statistics. For example, many statistics on private enterprises and establishments cover essentially the same universe of firms, yet each series employs its own coding method, sample selection methodology, etc, with the result that none of these statistics can be integrated to form a unified series. In other cases, series employ unique geographical grids, strata, or categories, which means cross referencing is often difficult and may lead to erroneous conclusions. The most well known

example is the apparent inconsistency in personal saving rates in the National Accounts and the Household Saving and Expenditure Survey.

Lack of coordination places a heavy burden on sample respondents, especially large firms that are included in most enterprise-based statistics. In 1993 more than 25% of polled firms listed on the Tokyo Stock Exchange said they had to reply to more than 100 different central and local government surveys each year[4].

Rectifying the situation is straightforward in some cases. For example, many establishment based surveys cover 100% of firms (their establishments) with more than 1 billion yen paid capital. It would be easy to use the same id code for these firms to facilitate cross referencing of a large variety of statistics.

2.3 Inadequate Disclosure

Inadequate disclosure of information is especially troublesome in two ways. First, many published statistics are processed using one or more primary statistical series but details of the procedure generally are not available. The disclosure problem is extremely severe for most of the National Accounts data, as they incorporate so many different statistics. (See Ando (2002) for the problems he faced in his exploration of the measurement errors in saving rate.)

In GDP statistics, the corporate sector includes not only private incorporated enterprises, but also the portion of activities of central and regional governments conducted by specific agencies (such as the postal system). There is no precise and reliable information on how to identify which part of the government activities are included. The problem is not limited to secondary statistics. The CPI is based on surveys of prices at sample retail stores, but original results are not available. As a result, it is not known (for example) how, and by how much, adjustment is made for quality change. The same problems exist for the WPI.

---

[4] The results cited above are taken from the following survey : Tokeichōsa Hōkoku to no kinyū ni kansuru jittai chōsa (Survey on the burden of respondents in official surveys and statistics), Sōmucho (to become Sōmusho in 2001).

For economists, an equally, if not more, important problem is government unwillingness to make original micro data available to outside researchers. The law explicitly and categorically prohibits use of official statistics for purposes other than the ones specified in the law establishing each statistical series or the corresponding ministerial orders. Thus, to obtain original data for designated statistics, one must file a petition for special exclusion. This is a complicated, time-consuming, drawn-out process with no guarantee permission will be granted. (See Matsuda et al (2000) for details.)

The difficulty in obtaining original data places severe constraints on outside observers, making it difficult even to point out with any reasonable accuracy where problems may be. Concern over the accuracy of CPI arose partly because many retail firms started publishing their own price data to argue that the CPI contains sizable upward bias (for example, Sezon Research Institute 2000). The resulting debate ultimately was unproductive in part because Sōmusho would not disclose data comparable to those covered by the retailers.

3 CPI Statistics

The consumer price index (CPI) in Japan is collected and published by Sōmusho Tōkeikyoku (Statistics Bureau and Center, Ministry of Public Management, Home Affairs, Posts and Telecommunications). Japan's CPI is by and large typical of CPIs collected in most countries. It is essentially a fixed-weight Laspeyres index, with weights taken from the Family Income and Expenditure Survey (FIES), which also is conducted by Sōmusho. The weights are revised every five years, incorporating the latest FIES.

Especially since the late 1990s when deflationary pressure became apparent, the CPI index has been criticized for its apparent failure to register the impact of rapidly declining retail

prices as reported in the media and by some of the largest national general merchandise stores (GMSs). [5]

Compared to the CPI in the United States, there are several notable differences in data collection procedures and lower-level aggregations. The Japanese CPI includes a larger number of individual items (roughly 600 compared to about 200 in the US). For each item, to survey prices, Japan uses a single brand and a single retail outlet within each designated area. Both outlets and items are rotated in the US.

Surveys are prices on specific days of each month rather than averages over period or brands, as in the US. Arithmetic means are used are used in every stage of aggregation, rather than geometric means. (The US converted to geometric for lower-level aggregation in January1999, as recommended by the 1996 Boskin Commission Report.)

3.1 Alternative Inflation Measures

If the CPI inflation rate is so problematic, why not use some other measures such as the GDP deflator or WPI? In fact, all of these are used to measure inflation, and many view the GDP deflator as a better indicator than the CPI. However, the same primary price survey data are used to estimate GDP deflators as to estimate CPI and WPI. So, if CPI and WPI data contain measurement errors, they will also appear in other processed statistics such as the GDP deflator.

Moreover, the CPI is a more appropriate measure of overall changes in the cost of living. In contrast, changes in the GDP deflator reflect overall changes in the prices of goods and services produced in the country, not necessarily those consumed. The difference can be large and important when events such as large increases in crude oil give rise to major swings in the final price.

---

[5] The most comprehensive study is Shiratsuka (1997). Shiratsuka (1999) offers in English a review of his 1997 monograph and other major studies. Sōmusho posts various documents prepared by the ministry on this issue at [www.stat.go.jp/data/cpi/8.htm.]

Table 3 shows the CPI, WPI for final consumption demand, and the GDP deflator for household final consumption. CPI and WPI are both Laspeyres indices with weights fixed for five-year periods, whereas the GDP deflator is a Paasche index with weights given by current-year expenditure shares. By construction, inflation in the GDP deflator has a downward bias, as opposed to an upward bias in CPI and WPI.

3.2 CPI as COLI, CPI as COGI

From the viewpoint of standard microeconomic theory, the principal objective of a CPI is to provide a benchmark for the cost of living index (COLI). However, as is the official view in most other countries, the Statistics Bureau of Sōmusho clearly states that the CPI should be viewed as the index of the specific basket of goods it contains – that is, the cost of goods index, COGI. It does not subscribe to the view that the CPI should be the best estimate of the cost of living index (COLI). (See Schultz (forthcoming) for a discussion of this incorporating the Boskin report.) Box 1 discusses the CPI as a COLI.

Box 1

CPI as a Cost of Living Index

Under certain strict conditions, we can derive a group of price indices, called Superlative Price Indices [see Diewart (1976) and Caves, Christensen, and Diewert (1982)] that approximate the true cost of living index up to the second order. One index among the group is the Tornqvist price index and it is given by

$$\log P_{0t}^{TR} = \sum_{i=1}^{n} \frac{1}{2}(\omega_0^i + \omega_t^i)(\log p_t^i - \log p_0^i) \qquad \text{EQ. 1}$$

where 0 denotes the reference period, i is the index for the goods and services, and ω is the expenditure share. The Laspeyres index, on the other hand, is given by

$$P_{0t}^{L} = \sum_{i=1}^{n} \omega_0^i \frac{p_t^i}{p_0^i} \qquad \text{EQ. 2}$$

The major advantage of a superlative price index, including the Tornqvist, is that the index properly incorporates substitutions among goods and services in response to, among other things, changes in relative prices. Neither Laspeyres (reference-period fixed weights) nor Paasche indices (current-period fixed weights) incorporate substitutions. The most serious problem with Laspeyres as an approximation of a COLI is that the index tends to over-represents prices that have risen from the reference period, thus over-stating the impact of price increases. By the same token, the index under-represents the impact of price declines. The magnitude of the bias depends crucially on two factors: relative prices and the degree of substitution across goods and services.

The practical difficulty in using Tornqvist or Fischer (geometric mean of Laspeyres and Paasche) indices is that they require current data on expenditure shares. If expenditure shares are continuously available, one can construct corresponding chained indices.

$$\log P_{0t}^{TRC} = \prod_{s=0}^{t-1} \sum_{i=1}^{n} \frac{1}{2} (\omega_s^i + \omega_{s+1}^i)(\log p_{s+1}^i - \log p_s^i) \qquad \text{EQ. 3}$$

$$P_{0t}^{LC} = \prod_{s=0}^{t-1} \sum_{i=1}^{n} \omega_s^i \frac{p_{s+1}^i}{p_s^i} \qquad \text{EQ. 4}$$

The important drawback of chained indices is path dependence. That is, the same magnitude of total price changes results in different price index values, depending on the sequence in which the changes take place. The problem is quantitatively important in high-frequency data. See Feenstra and Shapiro (2001) on such bias.

Even though we concur with the majority view among economists that CPI should serve as a measure of COLI, we also think that COGI, as it is constructed as an index representing a fixed basket of consumption goods, has its own merits. Especially as a

macroeonomic indicator, the inflation rate measured in terms of changes in COGI is important, given the crucial role played by the private and social costs of changing nominal prices. Unlike COGI, a properly defined COLI can change without any accompanying change in nominal prices, for example, due to changes in quality. This can be misleading especially when quality unadjusted indices are not available.

In relation to other price indices such as WPI, CSPI, and various wage indices, COGI is also important in monitoring the dynamics of vertical price formation. Thus, we agree that the CPI should continue to serve as a COGI, providing an aggregate measure of nominal price changes.

Even as a COGI, however, the CPI should perform better by incorporating lower-level substitution more explicitly: there is strong evidence that consumers substitute brands, shop around, and continue to shift toward mass retailers with lower prices. Moreover, unless one subscribes to an extremely narrow and rigid definition of a fixed basket (fixed brand purchased at fixed set of retailers), CPI should move in the direction of COLI at least in these dimensions.

We believe the CPI should serve both COLI and COGI purposes. Whenever important difference arises between the purposes, separate COLI and COGI series can be compiled. There is no practical or theoretical difficulty in this. As a matter of fact, the additional cost of preparing a separate COLI for different groups of household is relatively small, and the current CPI does include such series. We suspect, however, that the relevant COLIs for different groups differ substantially, once proper attention is paid to shopping behavior. To incorporate shopping behavior into the COLI, it is essential that information be collected at the household level.

Whether the CPI is viewed as strictly a COGI or also serves as a COLI, it is crucial to disclose details of the compilation processes, such as quality adjustments, as well as brand and sample-store replacements. Without full and timely disclosure of these details and the original survey results, the extent to which external monitoring can check potential problems is limited.

4 Major Sources of CPI Bias

There are several fairly well known, if not well-established, sources of problems in the Japanese CPI. All are considered sources of upward bias. One set relates to aggregation procedures and the second to lower-level data collection procedures (including how discounted prices are handled).

Collection procedures, services, quality change, and new products are covered in this section. Aggregation issues are taken up in later sections.

4.1 Lower-Level Data Collection Procedures

Under current procedures, prices for each item are collected first by specifying the most representative brand for each item, then by selecting the most representative sample store (usually the one with the largest sales volume of the item) within each precinct.

The brand selection procedure is problematic. Setting aside the problem of changes in the leading brand over time, fixing a particular brand in itself creates upward bias because many people are largely indifferent among brands and thus will substitute among brands, especially when one is temporarily discounted. Fixing a particular brand gives unbiased COLI data if and only if all consumers are completely brand loyal or retail prices of different brands all move together. Sōmusho does not release data on how many or how often brand replacements occur, but states that it checks the selection of specific brands every half year and replaces brands whenever appropriate.

In the United States, CPI does not fix any particular brand and different brands rotate in each price survey. The US procedure is superior in that the procedure avoids the inherent bias associated with fixing particular brands. On the other hand, Shapiro and Wilcox (1996) contend that brand turnover is closely related to CPI inflation in the US in that the bulk of the inflation rate is attributable to the imputed price increase registered for newly surveyed brands and entry-level (new) items when the sample is changed. That is, if brand A is substituted for brand B in the sample, the price difference between the two will be recorded as a price change affecting the CPI whether or not there is an actual change in the price of either brand between sample periods.

Selection of a single store within each sample precinct also is problematic, because consumers substitute among shopping outlets. Neglecting store substitution tends to introduce upward bias.

Discount prices (specials) are another issue. Each month the survey collects prices on the Wednesday, Thursday, or Friday of the week that includes the 12th. If the price is a discount price, the sample is void unless the price has been quoted for at least eight days at the time of the survey. It is not clear how regular and discount prices are defined. In most cases, the highest selling price seems to be the one defined as the regular price. It is unclear if the regular price ever changes at each store and, if so, how often. But, actual prices change quite frequently.

The current procedure thus tends to ignore almost all discount prices (whatever that means) of short duration. However, the bulk of sales of many products, especially ones easily stored, are concentrated in short periods when prices are discounted.

The extent to which discount sales are used differs systematically across items, brands, and types of retail outlets. Discounts are widespread and routinely used by national brands, whereas most generic commodities without strong brand recognition are rarely discounted. Discounts are far more common at large supermarkets and speciality stores, but very infrequent at small general stores and almost non-existent in convenience store chains.

Although there is no *a priori* reason to believe these measurement errors inherently generate systematic bias in the measured inflation rate, the recent macroeconomic setting and secular changes in the retail industry do give reasons to suspect they create systematic upward bias. The share of retail sales in Japan has been shifting away from traditional small stores toward large supermarkets and discount stores in suburbs, and toward inner-city convenience store chains. This may introduce systematic upward bias to the extent that current CPI procedures subsume some of the pure price differences across different types of stores as reflecting differences in service.

Biases created at lower levels can be quantitatively large precisely because they occur as a result of substitutions over very close substitutes: over time of the same brand, among different brands of the same good, and among neighborhood stores.

4.2 Services

After the revision in 2000, services comprise 48.4% of the CPI. There are no natural measures for the quantity of most services purchased. This implies that expenditure data such as FIES are ill-suited as the alternative data source for prices. Objective measurement of the quality of services is even more difficult. For these reasons, we have little to offer on biases from services.

Compared to commodity prices, there are reasons to believe raw price data are more accurate for some services in the CPI. For example, most utility rates and public transportation service prices are uniform and well documented. For these, there is little or none of the discounting so common for food and clothing. This applies also for price data on medical services. The bulk of payments are covered by public health insurance, and readily available and highly comprehensive price lists exist for individual treatments, various fees, and prescription drugs.

Setting aside quality issues, the biggest problem in service categories is under-representation of medical and health care in the CPI, as the weight is based on consumer out-of-pocket expenditure in the FIES, totally neglecting payments for medical insurance. According to the Survey on Medical Expenditure, in 1999 30.9 trillion yen (8.1% of national income) was spent on medical care. Out-of-pocket expenses covered by FIES were only 14.6% of that. In the current CPI, the weight for medical care is 2.4% and for health care is 1.4%, a total of 3.8%.

The medical- and health-related items in the CPI are limited to those not covered by typical health insurance. Thus, non-prescription drugs, physical check-ups, and the basic hospitalization fee for normal delivery of a baby are included, but most other medical services are excluded. Not surprisingly, data indicate systematic differences in price indices, depending on who directly pays the cost: the consumer, insurance, public institutions, etc. (See Iwamoto (2000) for some representative medical price indices).

It also should be noted that the CPI contains several conceptual flaws in some other service prices. Especially noteworthy is imputed rent for home owners. The actual rent data collected are those for rented dwellings; it is known, however, that rented and owner-occupied

homes differ greatly in capacity and quality. To the extent that the recent improvements in the quality of owner-occupied homes are not properly incorporated, measured rent is likely to include sizable upward bias. It should be borne in mind, however, that given the sheer magnitude of the diversity of dwellings across regions, types, and vintage, it is a formidable task even to estimate the size of the bias, let alone correct it.

4.3 Quality Change and New Products

Although quality changes and new goods are potentially the most important source of bias in the CPI, we do not investigate the problems in any depth here. Instead, two points. First, we argue that, in principle, CPI would benefit enormously from careful and systematic improvements in incorporation of the effects of quality change and introduction of new products. Second, there is an important inconsistency between CPI and WPI regarding certain groups of items. We suspect the inconsistency stems at least partially from differences in quality adjustments in the two indices. This is dealt with in a later section.

Some argue that, ultimately, measurement of quality should be aimed at measurement of contribution to the quality of life. For example, some say the measurement of medical services should be reformulated to measure the cost of cure, rather than the cost of treatment as is now the case. (See Schultze and Mackie forthcoming.) We do not engage in this debate here, except to the extent it is an aspect of the issue of the role of the CPI as a COLI as we noted earlier.

In the current CPI, essentially nothing is done to address the effect on living costs from introduction of the new products. This is understandable, given that no established procedure to do so exists. On the other hand, the long delay in incorporating changes in the consumption basket by itself introduces large and rectifiable biases if price declines primarily come soon after a product appears and before it is included in the CPI. That seems to be the regular pattern for many consumer durables, but it is conceivable that for other types of products, prices rise during the early stage.

It is only in the 2000 revision that the CPI included items such as personal computers and service charges for mobile telephones. The CPI still does not include fax machines, printers and other computer peripherals, or internet service-provider charges!

As for quality change, in the current procedure, whenever a sample item or brand is considered different in quality from the previous item, an overlap method is used to take account of quality changes. In 2000, the CPI for the first time started using hedonic methods to estimate quality changes in personal computers, but, as of now, this is the only item utilizing the method.

Few empirical studies in Japan measure quality changes and assess the impact of changes on the CPI. Shiratuska (1997, 1999) are the only published results we are aware of that estimate the impact of quality change on CPI bias. He estimates that under-estimates of quality changes result in an annual upward bias of 0.3% to 0.9%, with 0.7% the point estimate. However, he notes the estimate is based only on studies of a few consumer electronics and passenger cars.[6]

Most of Shiratuska's work uses data from the first half of the 1990s, so it is not clear if the same estimates apply to later periods. As shown below, in the late 1990s the consumer electronics component of the CPI registers a lower (actually, larger negative values) inflation rate than the comparable WPI rate.

For the United States CPI, Hausman (1999) estimates annual upward bias of 0.8% to 1.9% for telecom services as a result of not including cellular phone services in CPI until 1998. The potential bias can be substantially larger in Japan in that the use of mobile phone increased so fast and the price declined so dramatically. In 2001 the number of cellular users surpassed the number of fixed telephone lines in Japan.

Sōmusho (2000) has conducted preliminary estimation of a hedonic price index for personal computers. They estimate a price decline from the 1995 average, set at 100, to 12.8 by mid 1999. This is a 36.7% annual decline. Thus, if the personal computer had been included in the CPI in 1995, that alone would have reduced the inflation rate by 0.2 percentage points each year during 1995-99. (The personal computer weight in the current CPI is 0.54%). One can expect similar dramatic price decline for other items that now command sizable

[6] Shiratsuka (1997) and his associates estimated hedonic price indices for personal computers, camcorders, automobiles, and apparel. They found quality-adjusted personal computer prices declined 25% a year from 1990 to 1994, while unadjusted prices fell 3%. For camcorders, the annual quality-adjusted decline was 11%, but only 6% unadjusted. For automobiles, adjusted prices declined 0.4% annually, but increased 4% unadjusted.

expenditure shares: fax machines (not included), printers (not included), mobile phones (0.74%), internet service providers (not included), etc.

More often than not, the same goods and services appear on lists related to both proper adjustments in quality and timely inclusion of new goods. This is because the most important quality changes typically take place when items are relatively new. In this sense, timing is crucial. If an item is included only after it has become a part of the standard consumption basket, much of the impact of quality change and consumer surplus associated with quality-adjusted price declines is missed.

5 Aggregation Biases

Aggregation procedures are a problem. The Japanese CPI is a fixed-weight Laspeyres index. The biases created by using fixed weights and taking arithmetic means are well known. Aggregation bias arises at every stage in the Japanese CPI.

At the bottom level, one representative brand of each item is chosen for data collection. This assumes away inter-brand substitution and thus tends to create sizable upward bias. Fixed-weighting problems also appear in the selection of sample stores. As discussed later, this became serious in the 1990s as sales shifted away from small independent stores to larger chain-store discount outlets (see Table 8).

The FIES has a significant sampling problem in that it does not include single-person households. Given the large portion of the population living alone, and the substantial deviation of consumption patterns of single-person households from others, the bias implicit in this procedure is potentially important. Starting in late 2002, FIES is being expanded to cover single-member households.

In an earlier step to improve data quality, in October 2001 Sōmusho started a new consumption survey, covering 20,000 households, focusing on items the basic FIES is ill-suited to cover such as high-priced products purchased infrequently and services. Included are appliances, personal computers, other consumer electronics, mobile phones, and internet service providers, as well as some services already covered in FIES. The survey is conducted by a semi-private research organization. It includes single-member households.

Zero or negative inflation in recent years probably has lessened the size of aggregation bias in comparison with economies with a mild but positive inflation rate.

5.1 Higher-Level Aggregation Bias

At higher-level aggregation, it is well known that the current fixed-weight Laspeyres index using arithmetic means tends to produce some upward bias in the CPI. This is the case because whenever relative price changes, people do tend to buy more of the goods and services whose relative price declined, and buy less of those which have become more expensive, viz., in response to changes in relative prices, they change the consumption shares. The assumption of fix weights neglect this substitution and hence tends to overstate (understate) the impact of price increase (decrease). This problem of using fixed weights is not unique to the Japanese CPI. The procedure to measure the bias is simple and straightforward: annual expenditure weights from FIES for the 85 lowest-level categories are used to compute chained Fischer and Tornqvist indices, which are compared to the CPI, which uses the same price data but with fixed 1995 weights. Table 4 summarizes Shiratsuka's calculations and extends them to 1995-2000.

The bias is not large for years since 1995, except for 1999. Relatively large bias in the CPI inflation rate for 1999, i.e., the change in CPI from 1998 to 1999, probably reflects relatively large changes in consumption weights after the increase in consumption tax from 3 to 5% in April 1998. Compared to chained Tornqvist or Fischer indices, the fixed-weight Tornqvist generates roughly 0.054% upward bias per year in the five years through 2000. There is larger bias in earlier periods – on the order of 0.1%.

Although the magnitude is not large, aggregation bias is serious because it always exists and accumulates forever. Thus, it can have a quantitatively large impact when tracing living standards for generations.

Aggregation bias arises due to the under-representation of the scope of substitution whenever the relative prices of goods and services change over time. The results indicate that the bias is smaller in the more recent years primarily because of smaller variations in relative prices.

Notice that a low or negative inflation rate *per se* does not reduce aggregation bias. What matters is changes in relative prices. These results only confirm that relative price variability at higher-level aggregation is positively correlated with the inflation rate.

5.2 Discounts and Intertemporal Substitution

Biases created within each item, an aspect of lower-level aggregation, is now considered. There are two issues: selection of a particular brand of an item, and how price observations are collected. In a sense bias at this level is the easiest to deal with because, in principle, there is not much room for disagreement. The extent to which different brands of an item are substitutable is an empirical question that can be answered with reasonable accuracy if sufficient data are collected. Substitution across brands within each item is addressed later.

Here the issue is substitution over time of the same brand – that is, the extent to which consumers can exploit periodic discounts. This appears to be quantitatively important. How much depends primarily on consumer knowledge and the ability to hold inventory at home. (Feenstra and Shapiro 2000 is an early attempt to incorporate home storage and shopping patterns into CPI measurements. Also see Ariga, Matsui and Watanabe 2000.)

In principle, the upward bias due to the survey procedure described earlier applies only to the level, not necessarily to changes, in the index. The problem is essentially that the procedure systematically truncates the low price observations. This truncation may or may not generate upward bias in the inflation rate. Circumstantial evidence indicates, however, that it does indeed produce sizable upward bias in the measured inflation rate, as retailers reduce average sales price by further lowering the discounted price or increasing the frequency of discounts.

The easiest way to demonstrate the inflation bias created by intertemporal, intra-brand substitution is to compare the actual average purchase price to hypothetical price data, which CPI would collect following the data collection procedure described earlier. For this exercise, we use POS-DEI data. (See Appendix 1 for details regarding the data sets).

Table 5 covers six selected items sold at sample large-scale retail stores during the 24 months starting April 1995. The results are consistent across all of the items: namely, the

current CPI procedure consistently over-estimates the inflation rate because most special-sales prices are dropped from the survey. Notice that the results indicate that the decline in the average purchase price occurred primarily as a result of lowering the discount price or increasing the frequency of the discounts. Moreover, as Shiratsuka (1997) pointed out, the current procedure substantially increases noise, as it only sporadically picks up sales discounts. Table 5 shows that standard deviations in the inflation rate under the current survey procedure are substantially higher than those of average purchase prices for most items.

Unfortunately there is no unambiguous way to estimate the extent to which the bias due to survey procedures applies to other items in the CPI. It is known that periodic price discounts (specials, sales) are quite widespread in most medium- to large-scale retail stores. Discounts typically apply to processed food, toiletries, cosmetics, household appliances, and some clothing. In other words, for most items sold at large-scale retail stores, one expects periodic discounts. Table 5 indicates that the current CPI creates systematic upward biases for these items mostly in the order of 3% per year.

5.3 Substitution Across Brands

The CPI chooses a single brand to represent the price movement of each item. In general, ignoring substitutions across brands results in an upward bias in the level of the cost of living, but it is not certain if it results in any bias in the inflation rate. If the relative price of different brands is stable over time, the bias may well be negligible in computing the CPI.

Figure 1 shows three price indices compiled from POS-SRI data for liquid condiment, one of the 14 items included in the 1997 *National Survey of Prices, Special Volume on Bargain Prices*. Along with the Tornqvist index for the item, the figure shows indices for the brands that registered the lowest and the highest inflation rate from 1995 to 2000. Variations across brands are very large indeed.

Table 6 shows the intra-item sample variances for the 14 items and the monthly inflation rate for the corresponding item-level Tornqvist index. A simple panel regression of monthly item-level price variances on inflation rate for 14 items yields

$$\mathrm{var}_t^i = \sum_{k=1}^{14} const^k - .0363 ifr_t^i (.023) \qquad \text{EQ. 5}$$

$$R^2 = .0874$$

The result (standard errors in the parenthesis) indicates that deflation (-*ifr*) coincides with increase in price variations across brands. These findings thus indicate that, at least for these 14 items, consumers have ample opportunities to substitute among brands.

In Ariga, Matsui, and Watanabe (2000), we used daily POS data for two rival brands of curry paste sold at selected supermarket stores. Table 7 shows the impact of price discounts on sales volume.

Average sales volume of Brand A at a discount price is 57.4 if the Brand B also is at a discount price, which is 19% smaller than the average sales volume (70.7) at a discount price if the Brand B is sold at regular price. The impact of Brand A's discount on sales of Brand B at discount price is even larger, more than 30% (compare 29.2 against 42.4). On the other hand, pricing has a much smaller impact on volume at regular price, around 6% to 8% (5.0 versus 5.4 for Brand A, and 2.9 versus 3.1 for Brand B).

Given the large impact of periodic price discount on sales, these figures suggest the presence of heterogenous consumers, as well as sizable inter-brand substitutions in response to changes in relative prices. Although these findings strongly indicate that price data of any particular brand can be a highly misleading indicator for overall changes in prices of different brands of each item, it is not possible to provide estimates of the magnitude of the inflation rate bias created by brand substitutions *per se*. Given the analysis on inter-temporal substitution, it is probably not very productive to try to estimate the effects alone, as substitution in this aspect is closely related to inter-temporal substitution and periodic price discounts.

There also are difficulties from the extremely high rate of new brand introductions and retirement of old brands, particularly among items in the food, household appliances, toiletry, and clothing groups. Shifts in sales shares from one brand to another not only are highly frequent but also unpredictable. This makes it practically impossible to obtain reliable estimates of substitution elasticities for the wide range of goods in the CPI. Again, these observations indicate the problem inherent in choosing a single specific brand to represent the

spectrum of brands of each item. It is far more satisfactory and actually easier to use price averages across brands.

5.4 Substitution Across Stores

According to the current CPI procedure, the survey selects the most representative store within each survey precinct for each item. Nation-wide, the survey has roughly 700 precincts. Usually the store with the largest sales volume is chosen for the item.

Table 8 shows the changes in shopping points in the National Survey of Family Income and Expenditure. As expected, regular stores lost shares across the board in the 15 years 1984-99. The decline is especially large in food.

Sōmusho (2000) explains the selection procedure for precincts and sample stores. It is not entirely clear, however, to what extent the delay or failure in changing sample retail stores contributes to selection bias in the CPI. According to Sōmusho, "The latest store selection is fairly close to the 1999 distribution" shown in Table 8.

Shiratsuka points out that "the shift from department stores and general small stores to discount outlets has largely subsided," so that price differentials have "settled down to a level consistent with the difference in service quality" (1999, p 90). However, Table 8 suggests the shift is still very much an ongoing process.

The current CPI revises sample store selections in two ways. Every five years the most representative store is chosen for each commodity group in each precinct. This reflects changes in market shares across different types of retail shops in each precinct and commodity group. In principle, the CPI uses the overlap method to correct for underlying differences in retail services between sample stores before and after the changes. Sample stores also are replaced on an ad hoc basis. This is necessary when stores are closed or stop selling the sample product. In such cases, price data are directly connected and no adjustments are made in prices. In the case of services, the overlap method is used.

To sum up, the current procedure uses direct-comparison methods only for ad hoc sample-store replacements for commodities. One expects that in the case of an ad hoc replacement, the replacing store is selected in a way that retains the characteristics of the

previous sample store. It is not clear to what extent overall the overlap and direct-comparison methods are used. As a result, it is not known how much of the price differentials across stores are subsumed and assumed away using the overlap method. We suspect that whenever major changes in the characteristics of sample stores occur, the overlap method is used so that the CPI attributes the price differentials across old and new sample stores to differences in the quality of retail services. In short, even if the CPI has been correctly adjusting the sample store distribution to changing shopping patterns, most within-brand price differentials across different types of stores are assumed away.

In principle, we agree that some price differentials reflect differences in service quality. On the other hand, given the long history of restrictions on entry of large-scale retail stores, and the fact consumers do shift purchases from general small-scale stores to supermarkets and mass-marketing specialty stores whenever such stores are opened in the neighborhood, it seems clear that some of the price differentials are indeed pure price differentials, reflecting the local monopoly power element of retail pricing. Table 9 and Table 10 offer some evidence, using cross-section data on retail prices of 14 items at a variety of retail stores at many locations collected by the 1997 National Survey of Prices.

Table 9 shows the difference in actual retail prices of the items across different types of stores. Ariga, Matsui and Watanabe (2000) found that for two brands of curry pastes sold at sample supermarkets, 31% of daily observations were of discounted price, but 72% of volume was sold at discount prices. More generally, for a sample 18 supermarkets we found 70% as the share of sales at discount prices. General small stores and coops offer price discounts much less frequently.

To indicate that some of these price differences reflect pure price differences, we used the survey data to run simple cross-section regressions on average regular and sales prices over a set of dummy variables, including one representing the presence of nearby rival stores. The results in Table 10 show that both regular and discount prices are significantly lower among stores with nearby rival stores.

Specifically, among regular small-scale stores, the regular price is 8.2% lower than comparable stores without a nearby rival. The impact of a nearby rival on the discounted price is 15.5%. In other words, the results suggest that a significant portion of price differences

between large-scale and small-scale stores reflects the effect of local competition on pricing, rather than differences in service quality.

The same source shows that 26% of small-scale regular stores reported no nearby rival, whereas for large-scale supermarkets, only 3.7% reported no nearby rival. Notice also that the impact of a nearby rival on prices is far smaller in the case of supermarkets, mass-marketing speciality stores, and coops. Setting aside the difference in geographical sizes of markets for respective types of stores, the data indicates strongly the presence of monopolistic power of many small-scale retailers.

We conclude from these results that sizable price differences exist between small-scale general retailers and large stores, and that some of these differences reflect lack of local competition for some small-scale retailers.

As indicated in Table 8, continuing shifts in sales share away from small-scale to large-scale stores should have generated sizable price declines for average consumers. For the sake of argument, suppose on average that a 10% pure price difference exists between the two types of retailers. This implies a roughly 0.1% upward bias in the CPI from not accounting for the pure price differences resulting from shifting shares. This is computed by multiplying the 6.5% decline in the share of small-scale stores by the 10% price differential over 5 years. In any case, unless we know the extent to which the overlap method is used for each type of sample store replacement, the effect on CPI bias cannot be estimated with any degree of accuracy.

The current store selection method poses other problems. The price differences in Table 9 are likely to generate sizable variations in average purchase prices across households, depending on residence location, income, member composition, age, and other attributes. Choice of a single representative store in each precinct for each item inevitably masks these variations. Such considerations are important if the CPI is used as a COLI. More generally, the current CPI system is ill-suited for incorporating cross-sectional and inter-temporal variations in shopping behavior, and this has consequences on the COLI.

6 A Curious Discrepancy between CPI and WPI

This section compares CPI and WPI data for two groups of commodities to get some idea on the likely magnitude of the bias created by quality change.

Until the mid 1990s, with the exception of consumer electronics, the CPI inflation rate tended to be higher than the WPI rate for most items common to both indexes. Circumstantial evidence suggests significant upward bias in CPI or downward bias in WPI (or both) due to quality changes in the longer run, but at least since the mid 1990s this may not be the case. In the last ten years, the annual impact of all quality change on the WPI is estimated to be around 0.3-0.4% by the Bank of Japan (2001a).

The groups being compared are processed food and consumer electronics. The likely magnitude of quality improvement in processed food in the WPI is around 0.1% per year (Bank of Japan 2001a). Given the magnitude of the estimation error, we take the effect as essentially zero, and this is the primary reason the group is used in the analysis as the benchmark. For consumer electronics, the potential impact of quality change on CPI bias is one of the largest among items in the index.[7]

Using CPI weights, the average inflation rates of the two indices for the two groups using only items commonly found in both is shown in Table 11.

The result for consumer electronics implies retail prices declined relative to wholesale prices by as much as 25% during the 1990s. If the sample period is extended back to 1980, the average annual difference is 1.9 percentage points, which translates into a decline in relative retail price of as much as 66.4%. This is suspect because the distribution margin is at most around 30% of the retail price and available statistics suggest at most a modest decline in the retail margin during the period – perhaps a few percentages of the retail price. In other words, either CPI or WPI, or both must contain sizable biases.

One possibility is that WPI severely under-represents the price declines. In the 1990s, many consumer electronics firms relocated plants to Asian developing economies and the

[7] Automobiles have the largest effect: -3.1% per year on its sub-index. However, the WPI has indices for three different types of passenger cars, while the CPI has only one. Hence we decided to use consumer electronics as an example.

import of these goods quickly replaced domestic production. In the 1995 revision of WPI, the Bank of Japan started collecting import price indices of these products.

The bottom row of Table 11 shows the weighted inflation rate of consumer electronics during 1995-2000, with WPI replaced by the corresponding import price index. The result is essentially the same. Although the coverage of imported price indices is far from exhaustive, it seems unlikely that the deviation can be due solely to the rapid price decline of imports. Another possibility is that the large difference in price levels between domestic and imported products is the root cause. The rapid decline of retail prices could reflect rapid replacement of high-priced domestic items by cheaper imports even if the imported goods' prices did not decline faster than the domestic ones.

It is conceivable that the Bank of Japan has severely underestimated the underlying quality changes of these products, more so than Sōmusho did for the CPI. We consider this highly unlikely, given the nature of the debate between Bank of Japan and Sōmusho on the possible upward bias of CPI. Another possibility is that CPI over-estimates quality change and so under-estimates the inflation rate for this group. There is reason to believe that hypothesis has merit and it thus needs further investigation.

There are differences in quality adjustment methods between the two indices. According to Bank of Japan (2001), the most popular method for dealing with quality change in the WPI is cost comparison. It is used for about 30% of WPI items. In contrast, Sōmusho (2000) states that the CPI uses either the overlap or the direct comparison method. Although Sōmusho does not reveal how many items are quality adjusted by which methods, it says that "whenever a sample brand is replaced, unless there are reasons to believe that the new and old brands are essentially the same quality, the overlap method is used" (author's translation). Hence, it is reasonable to say that virtually all substantive quality adjustment in CPI is done using the overlap method. Bank of Japan also uses the overlap method, but only on about 10% of WPI items.

Overlap methods can generate sizable over-estimate of quality change if the retail price of the existing brand declines substantially in anticipation of a forthcoming future brand. Suppose the CPI survey collects prices for brand b until period t and then replaces it with b' at $t+1$. Replacement typically occurs because of a decline in the brand's market share or its

disappearance from the sample store. Overlap methods treat the price differential between the current and replacement brand b' as reflecting an underlying quality difference, so the price index for item i is computed as

$$p^i_t = \frac{rp^b_t}{rp^b_0}$$
$$p^i_{t+1} = rp^{b'}_{t+1} \frac{rp^b_t}{rp^{b'}_t} \qquad \text{EQ. 6}$$

where rp^b is the survey price of a particular brand. Substantial over-estimate of the quality change can occur if the relative price

$$\frac{rp^b_t}{rp^{b'}_t} \qquad \text{EQ. 7}$$

does not properly represent the quality difference. In particular, a disappearing brand might be heavily discounted around the time of replacement. In that case, quality improvement is over-estimated and the method introduces downward bias in the inflation rate.

Sōmusho (2000), using color televisions as an example, reports that a chained index using overlap methods generates a 46% decline in the index for the three-year period 1995-98, which can be compared to a decline of 27% in the hedonic price index and 25% in the published CPI index. On the other hand, estimates by Shiratsuka (1997), discussed earlier, suggests significant upward bias in CPI due to under-estimation of quality change during the first half of 1990s. Our results shown above cast some doubt on the alleged upward bias in CPI for this reason.

All in all, for the late 1990s, we cannot make any definitive statement on even the direction of bias created by quality change. But, in any case, it is certain that there are important inconsistencies in quality adjustments between the CPI and WPI for at least some product groups.

7 Impact of Shopping Patterns on COLI

The current CPI almost totally ignores the impact on COLI of diverse shopping patterns by different types of consumers. This is also true of CPIs in most other countries. In Japan, there are supplementary CPI indices incorporating differences in consumption patterns across different types of households. They do not incorporate the impact of shopping patterns on respective COLI, however.

In Appendix 2 we develop a simple model of cost minimization and demonstrate the impact of shopping and storage costs on shopping and purchase decisions. Two points emerge. First, pricing patterns of retail stores significantly influence consumer decisions on shopping timing and purchase. Second, large variations in shopping and storage costs, as well as average purchase price, result from variations in pricing policy across different types of stores. Moreover, variations in consumer shopping and storage costs influence which store is the optimal choice. These results suggest that the variation in COLI across regions and household types can be much larger than what the current CPI indicates.

8 Estimation of Commodity CPI Biases

Inevitably, estimation of bias involves many subjective judgments and is likely to contain sizable errors. The potential impact of each source of bias differs across categories, as does our ability to estimate its direction and magnitude. For this reason, our analysis on bias will be confined to the commodity CPI. Services are not considered.

Commodity CPI comprises 51% of overall CPI. We provide two results. The first compares CPI with COLI using unit prices in FIES. The second is the COLI for 14 selected items using POS-SRI. The two are consistent in suggesting sizable upward bias in commodity CPI.

Table 12 compares four COLI indices for a variety of CPI categories. In the comparison, unit price indices in FIES are used because CPI item selection is based on FIES, which collects unit prices for about 200 items.

The large deviation between the two indices for clothing (2% per year) is consistent with consumers rapidly shifting from domestic to imported, and from small-scale to

mass-marketing speciality stores. This shift started with the rapid expansion of several chain stores specializing in men's suits and other formal clothing. The department stores, traditionally the most popular choice for such items, lost share. Beginning in the late 1990s the shift has been concentrated in more casual clothing and underwear. Among others, the UNIQLO chain registered explosive growth in sales and profits.

Table 13, comparing POS-SRI Data with the CPI shows an upward bias in the CPI on the order of 1.5% per year. For 6 of 14 items selected in the 1997 *National Survey of Prices, Special Volume on Bargain Prices*, FIES also reports unit prices. The difference from the CPI for these groups is again around 1.5% per year. These estimates are very close to the bias estimated in Table 4. Although the two baskets differ, an index computed by aggregating all FIES items yields a 1.35% lower inflation rate than the overall CPI.

The conclusion is that, for at least food and clothing groups, the CPI since the mid 1990s has sizable upward bias, most likely in the range of 1.5% to 2% per year. We believe a bias of similar magnitude exists for other items commonly sold at mass retail stores (such as appliances and toiletry goods), so that all-told roughly two-thirds of commodity CPI belongs to groups we believe are biased upward by 1.5-2.0% per year.

To be conservative, assume the bias arises only for purchases of these commodities at large retailers and that two-thirds of purchases are at mass retailers. Applying the low end of the bias range, 1.5% per year, suggests a bias of 0.67% in the CPI. Using 2.0%, the impact on CPI is roughly 0.9%. Even assuming the CPI bias is zero for other commodities and also for samples taken at small-scale stores, the effect on overall commodity CPI must be 0.5% to 1.0% per year. The difference between unit price inflation in FIES and the CPI inflation rate among comparable items other than consumer electronics and services is about 0.6% per year (Table 12), which is within the range just estimated.

We believe that 0.5% to 0.6% per year is a conservative estimate of the upward bias in the CPI as a measure of COLI because service prices, which comprises roughly 50% of the overall CPI, have not been covered in the analysis. Upward biases in many important items in this category is likely. On the other hand, the comparison of CPI with WPI indicates a potential downward bias in the CPI.

9 Some Suggested Ways to Improve CPI

Japan's CPI contains upward biases and has other problems. Some of the problems can be corrected or at least alleviated. Here are some suggestions for improving the CPI.

9.1 Upgrade Statistics Sections

The Statistics Bureau of Sōmusho, and most other statistics sections of Japan's central government, are seriously under-manned and suffer from meager budget allocations. There are fewer highly trained statisticians than is appropriate for the work, and there are no staff members with advanced economics degrees. Not only must more people be hired, but the new hires should be specifically skilled.

Staff and budget constraints severely limit the options available to improve CPI. For example, use of POS data is highly expensive as Sōmusho has to purchase them from the outside private sector. Needless to say, collecting POS data by themselves is even costlier and practically impossible. Systematic attempts to estimate hedonic price indices require large resources for data collection and estimation. In the United States, the Bureau of Labor Statistics (which prepares the US CPI) quickly incorporated recommendations in the Boskin Commission report on problems with the US CPI (Schultz and Mackie forthcoming). Given the budget and staff size limitations, it seems very difficult for Sōmusho to carry out similar research with comparable speed.

9.2 Improve Data Collection

Data collection methods need to be changed in many aspects, most of them fundamentally. First, the revision of item selection and weights must be done more frequently. In principle, to the extent the CPI uses FIES, this is a matter of automatic adjustments. FIES is monthly, but annual CPI revision is a more realistic goal. The need for continuity can easily be met by tracking CPI component indices based on weights and item selections in the past. The additional tasks created by annual revision may not be large.

Utilizing other official data sources in compiling the CPI offers significant benefits. For example, the gain from coordinating data collection and compilation for CPI and WPI is obvious. Coordinating with other agencies also should be done, especially regarding service prices. In particular, there should be large gains in accuracy from utilizing other sources of data on medical and health care, and housing expenses.

A more fundamental change is to seek alternative data sources. Current collection relies exclusively on surveying sample retail firms. Given the time and resource constraint, the margin of improving data quality in commodity CPI may be fairly narrow to the extent the current method is retained. However, we propose two alternative (complementary) data methods.

The first is to use POS data, which is available on a daily basis for essentially all the brands sold in sample retail stores. Moreover, POS data contain quantity data totally missing in the current survey. Such data are important for several reasons. Even if Sōmusho retains its current position that the CPI should be based on representative brands, POS data provide more accurate and timely information on which brand is the most popular. Being available on a daily basis makes allowing for sales and temporary price mark-downs easy and straightforward. Sōmusho uses POS for collecting price information on one item – personal computers, since 2000.

The second complementary data source is to improve and modify FIES to make it usable as a source of CPI price information. The advantages of using consumer-side information are numerous. The consistency between the CPI basket and the actual consumption basket would be improved greatly. For the purpose of COLI, the actual mix of brands within each item and expenditure shares of items are the ideal set of information. To the extent FIES accurately represents these choices, there should be no disagreement on how to best represent the consumption basket and relevant purchase prices. Improving the selection of sample retail outlets will not be necessary, as consumers themselves make the choice, which can be observed.

Adjustments to incorporate quality change are the most difficult and this paper has not covered the issue in any detail. We are sure there are important inconsistencies between CPI

and other price data, especially WPI. The discrepancies are quantitatively large. Both CPI and WPI will benefit from proper coordination and joint work by Sōmusho and Bank of Japan.

9.3 Create an Independent Research and Appraisal Body

Resources should be used to establish an independent body to conduct research and systematic appraisal of major statistics. Such research is especially important for statistics compiled from many primary statistics, such as the National Accounts. Given the current state of information disclosure, and the inevitable information advantage of inside staff, such research must be conducted within the government, rather than completely out-sourced, although the research would benefit from using outside consultants.

The Statistics Council is a committee overseeing statistics collection and compilation activities of the central government. Although the council in the past made important policy recommendations to improve the official statistics, its abilities are limited. Like other government councils, members are non-government officials and meet only a few times a year. Without a body of research staff working on a regular basis to monitor official statistics, its recommendations are necessarily abstract in nature and often too late. Given the autonomy of individual ministries, it is unclear to what extent the council has influence on changes in individual statistics produced in different ministries.

10 Conclusion

We have employed a variety of data and alternative aggregation and estimation methods to estimate biases in Japan's CPI. The results strongly suggest the presence of sizable upward bias in the commodity CPI. Our best estimate is at least 0.5% per year excluding biases in services and from quality changes. The true bias is likely to be larger than this estimate, but far more extensive research is needed to obtain a more reliable figure.

After a journey into a maze of price data, we come back yet again to one of our first points: the Japanese government should allocate far more resources to collection, compilation, and timely disclosure of statistics. Although private data collection services have grown rapidly since the late 1980s, the need for official statistics is obvious and compelling. No

private sector entity can realistically replace the statistics collection activities of the central government.

The potential benefit from improvement in indices such as CPI can be enormous, given that so much decision-making is linked explicitly or implicitly to the CPI. Although many suggestions for improvements can be implemented within the current budget and staff allocations, the more fundamental necessary changes require sizable increases in budget and staff.

We have pointed out several times the need for coordination within the government. This is straightforward. Statistics based on the same population of samples should use compatible data strata, the same method for coding, and the actual surveys should be merged to the maximum extent possible in order to minimize costs to respondents. And, there needs to be an independent body within the government conducting research and appraising the statistics.

Although focused here on data collection and lower-level aggregation issues in CPI mis-measurement, we concur with the majority that problems associated with quality adjustments and introduction of new goods are by far the most important and challenging. Moreover, shopping behavior and retail competition needs to be incorporated into CPI. These and other issues are left for future research. The central message of this paper is the need for fundamental changes in the way CPI is collected and compiled.

References

Ando, Albert (2002), "Missing Household Saving and Valuation of Corporations: Inquiry into Japanese National Accounts," *Journal of the Japanese and International Economies* 16 (2): 147-176

Ariga, Kenn, Kenji Matsui, and Makoto Watanabe (2000), "Hot and Spicy: ups and downs in price floor and ceiling of Japanese Supermarkets," mimeo.

Bank of Japan (2001a), "Oroshiuri Bukka shisū no Hinshitsu Chōsei wo megutte" (On Quality Adjustments in Wholesale Price Index), a paper presented at a conference on Price Index.

Bank of Japan (2001b), "Oroshiuri Bukka Shisū ni okeru Hedonic Approach" (Hedonic approaches in the wholesale price index), Bank of Japan Working Paper 01-24

Boskin, Michael.J., Ellen Dulberger, Robert J. Gordon, Zvi Griliches, and Dale W. Jorgenson (1996), *Final Report of the Advisory Commission to Study Consumer Price Index*, U.S. Government Printing Office

Caves, David, W., Lawrence R. Christensen, and W. Erwin Diewart (1982), "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity," *Econometrica* 50(11): 1393-1414

Diewart, W. Erwin (1976), "Exact and Superlative Index Numbers," *Journal of Econometrics* 4: 115-144

Feenstra, Richard C., and Matthew D. Shapiro (2001), "High Frequency Substitutions and the Measurements of Price Indexes," paper prepared for the CRIW Conference on "Scanner Data and Price Indexes."

Gordon, Robert J. (1999), "The Boskin Commission Report and Its Aftermath," *Monetary and Economic Studies* (Bank of Japan), December 1999: 41-68

Hausman, Jerry A. (1999), "Cellular Telephone, New Products, and the CPI," *Journal of Business and Economic Statistics* 17(1): 188-194

Iwamoto, Yasushi (2000), "Jinkō-Kōreika to Iryōhi" (Aging Japanese Population and Medical Expenditure) mimeo.

MacDonald, James M., and Paul E. Nelson, Jr.(1991), "Do the poor still pay more? Food Price variations in large metropolitan areas," *Journal of Urban Economics* 30(3): 344-349

Matsuda, Yoshiro, Keiro Hamasuna, and Hiromi Mori (eds.) (2000), *Kōza Micro Tōkei*

Bunseki (Microscopic Statistical Analysis) in 3 volumes, Tōkyō: Nihon Hyōronsha

Matsuoka, Mikihiro, and Brian Rose (1994), *The DIR Guide to Japanese Economic Statistics*, Oxford University Press

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

Schultz, Charles, and Christopher Mackie (eds.), *At What Price? Conceptualization and Measuring Cost-of-Living and Price Indexes*, National Academy of Sciences, forthcoming

Sezon Research Institute (2000), "Ōte Ryōhan-ten no POS data wo riyō shita bukka shisū ni kakakwaru kenkyū," (Price Index using POS data of a Large Scale Retailors)

Shiratsuka, Shigenori (1997), *Bukka no Keizai Bunseki (Economic Analysis of Inflation Measures)*, University of Tokyo Press (in Japanese)

Shiratsuka, Shigenori (1999), "Measurement Errors in the Japanese Consumer Price Index," *Monetary and Economic Studies* (Bank of Japan), December 1999: 69-102

Sōmusho (Ministry of Public Management, Home Affairs, Posts and Telecommunications) (2000) *Shōhisha Bukka Shisū Sankō Shiryō #17* (Reference documents for CPI #17)

Table 1

Major Official Statistical Series, 2000

Designated[1]	Approved[2]	Budget (1999)[3]	Staff	Ministry
8	119	13,032	[a]5,979	Agriculture
4	50	256	102	Education
2	8	144	86	Finance
8	102	5,758	465	Health and Labor
7	68	4,169	124	Land and Transport
14	59	14,494	1,617	Public Management (Sōmusho)
17	47	5,867	381	Trade and Industry
-	27	1,360	50	Others
62	480	45,080	[b]8,804	Total

1 We include only those designated statistical series which are currently collected on a periodic basis, thus excluding those for which new data collection has been stopped. In effect, the latter series are no longer used, primarily because of the lack of interest (they retain the special status only because the use of the original data is still tightly controlled by law).
2 The number of approved series collected in each year at each ministry varies widely, but the total number has been stable between 400 and 500 since the mid 1990s. Unlike the designated series, many of these statistics are collected once and only.
3 In million yen for fiscal 1999, which ended 31 Mar 2000. This is roughly 0.06% of the central government budget. The budget has been in a 40-50 billion yen range since the early 1990s except when there is a population census (years ending in 0 and 5). Thus, the total fiscal 2000 budget was 98.6 billion yen, with 75.9 billion allocated to Sōmusho, which conducts the census.

a This is 68% of the total. Most of them are at regional offices of the ministry.
b The total given is 2.2% of total central government administrative staff, 398,000.

Source: *Tōkei Kijyun Nenpo (Statistics Standards Annual)*: Sōmusho, 1999, 2000.

Table 2

US Statistical Staff and Budget

Budget (2002)[1]	Permanent Staff Total	Permanent Staff Statisticians	
366.6	1,595	33	Agriculture
143.1	4,154	1,403	Commerce, except Census
563.4	3,708	1,398	Census Bureau
198.0	127	78	Education
1,260.6	606	212	Health, HS
57.4	67	42	Justice
655.4	2,792	179	Labor
122.3	162	54	Transportation
686.3	374	55	Other
4,110.5	9,877	2,056	Total
3,906.3	6,169	658	Total (ex Census)[2]

1 Estimate for Fiscal year ending 30 Sep 2003, in million dollars.

2 Excludes the 2000 census, entities spending less than $0.5 million, and statistics collection in conjunction with other major activities.

Source: US Executive Office of the President, Office of Management and Budget, *Statistical Programs of the United States Government*, 2002.

Table 3

Annual Inflation Rates, 1900-2000

(percent)

Period	CPI	WPI[1]	GDP[2]
1990-2000	1.64	-0.55	0.49
1995-2000	0.30	-0.76	-0.32

1 For final consumption goods

2 Deflator for household final consumption. This is a Paasche index using current weights from FIES. Both factors tend to generate a lower inflation rate than the CPI.

Table 4

Aggregation Bias

(annual percentage rates except as noted)

		Fixed Weights			Chained			
		Laspeyres	Tornqvist		Tornqvist		Fischer	
	CPI Index	Inflation	Inflation	Difference[1]	Inflation	Difference[1]	Inflation	Difference[1]
1996	100.180	0.180	0.154	0.026	0.154	0.026	0.154	0.026
1997	101.869	1.689	1.651	0.038	1.657	0.032	1.649	0.032
1998	102.613	0.744	0.713	0.031	0.716	0.028	0.703	0.028
1999	102.242	-0.371	-0.498	0.127	-0.440	0.069	-0.429	0.058
2000	101.415	-0.827	-0.877	0.050	-0.864	0.037	-0.870	0.043
1995-2000	-	0.283	0.229	0.054	0.245	0.038	0.241	0.042
1970-95	-	4.438	-	-	4.313	0.125	4.216	0.222
1990-95	-	1.153	-	-	1.152	0.001	1.272	-0.119

1 Difference between the inflation rate in the CPI in column 2 and the inflation rate in column to the left. As discussed in the text, this is an indicator of upward bias in the CPI inflation rate.

Source: 1970-95 and 1990-95 are from Shiratsuka (1998).

Table 5

Bias Due to CPI Data Collection Procedure

Mean Inflation [1]			Standard Deviation[2]		
Weighted Average	CPI Procedure	Upward Bias	Weighted Average	CPI Procedure	
-1.36	-0.38	0.98	.0148	.0208	Mayonnaise
-3.12	-0.41	2.71	.0195	.0305	Ketchup
-2.25	0.00	2.25	.0343	.0434	Soy sauce
-2.94	-0.30	2.64	.0238	.0429	Liquid soup base
-2.73	-0.10	2.83	.0298	.0149	Laundry Detergent
-5.44	-1.45	3.99	.0378	.0957	Instant Coffee

1 % per year

2 Annual log differences.

Table 6

Mean Inflation Rates and Within-Item Variances

Mean inflation rate(%)	Variance[1]	
0.4477	.016	Instant Coffee
0.1006	.0062	Facial tissue
-0.2258	.0041	Mayonnaise
-0.1324	$.94*10^{-5}$	Yogurt
0.0781	.0080	Liquid condiments
0.2465	.092	Fruit juice
-1.4649	.019	Fresh milk
-11.8878	.038	Sugar
0.2117	.0061	Wheat flour
-0.1937	.0043	Soy sauce
0.2850	.0070	Cooking Oil
0.9755	.052	Sanitary napkins
-0.0627	.060	Laundry detergent
-0.0738	.0078	Kitchen detergent

1 Monthly average for indices of brand-specific inflation normalized to set the annual average for 2000 equal to 1. The variance of mean inflation rate across different brands within each item is shown in the second column.

Table 7

Substitution Across Brands: Curry Pastes

If A's price is:		
Regular	Discounted	
		Brand A sales volume when B's price is:
5.4	70.7	Regular
5.0	57.4	Discounted
		Brand B sales volume when B's price is:
3.1	2.9	Regular
42.4	29.2	Discounted

Unit: average number of sales unit per day

Table 8

Share of Expenditures on Selected Items, by Type of Retail Outlet, 1984-94

(percent)

	Regular small-scale	Super-markets	Depart-ment stores	Conveni-ence stores	Coopera-tives	Dis-count[1]
Total						
1984	50.8	28.9	10.0	-	4.6	-
1994	40.5	30.3	9.3	1.1	6.3	4.0
1999	34.0	35.3	9.1	1.7	5.9	5.4
Food						
1984	40.6	44.9	3.5	-	7.0	-
1994	25.2	49.2	4.1	1.9	10.4	2.3
1999	16.7	57.5	4.5	2.6	9.8	2.7
Appliances						
1984	46.3	24.1	15.1	-	4.3	-
1994	37.1	22.7	10.8	3.1	5.8	12.5
1999	32.3	26.6	9.4	3.0	5.3	16.7
Clothing						
1984	36.9	18.1	37.6	-	1.8	-
1994	33.7	17.7	34.0	1.3	2.1	11.2
1999	28.7	20.5	36.5	1.4	2.3	10.6

Row totals do not add to 100% because not all store types are included.

1 Mass marketing speciality discount stores.

Source: *National Survey of Family Income and Expenditure*

Table 9

Average Across-Store Price Differentials for 14 Items

Small-store regular price = 100.

Small stores	Super-markets		Mass discount		Coops		
100	95.4		95.0		94.4		Regular Price
78.7	64.9	(17.5%)	68.9	(12.5%)	68.7	(12.7%)	Discount Price
89.4	74.0	(16.7%)	76.8	(14.1%)	81.6	(8.7%)	Case 1[1]
95.7	74.0	(22.7%)	76.8	(19.8%)	89.3	(6.7%)	Case 2[2]

Percentages in parentheses are the discount from the small store's price for each of the cases.

1 In determining the average price for each category, 50% of volume is assumed to be sold at a discount.

2 In determining the average price for each category, 20% of volume is assumed to be sold at a discount.

Data source: 1997 *National Survey of Prices Special Volume on Bargain Prices*.

Table 10

Impact of a Nearby Rival Store on Retail Prices

log (Regular Price)		log (Discount Price)		
-.0401	(6.16)	-.0426	(4.50)	Large Store
-.0579	(2.84)	-.243	(8.18)	Supermarkets
-.170	(4.24)	-.315	(5.42)	Mass-discount
-.116	(2.63)	-.180	(2.79)	Coop
-.082	(4.48)	-.155	(5.82)	Rival Store (RS)
.0659	(3.14)	.122	(3.98)	RS*Supermarkets
.0898	(2.17)	.139	(2.31)	RS*Mass-discount
.120	(2.62)	.123	(1.86)	RS*coop
.995		.978		Adjusted R^2

Results of OLS cross-section regressions.

Numbers in parentheses are t-statistics.

Data source: 1997 *National Survey of Prices Special Volume on Bargain Prices*.

Table 11

Comparison of CPI and WPI

(annual percentage rates)

1980-2000		1990-2000		
CPI	WPI	CPI	WPI	
0.83	0.68	0.25	-0.35	Food
-3.33	-1.49	-6.12	-3.32	Consumer Electronics
-	-	-5.76	-2.49	Import Price Index

CPI weights are used for both CPI and WPI

Table 12

COLI Average Annual Inflation in CPI and FIES

(percent)

1980-2000

(C,C)[1]	(C,K)[2]	(K,C)[3]	(K,K)[4]	CPI-FIES[5]	
0.83	0.70	0.44	0.56	0.27	Food
1.31	1.41	0.20	0.03	1.28	Clothing[6]
-8.62	-6.05	-4.80	-3.41	-5.21	Consumer Electronics
0.11	0.45	-0.55	-0.69	0.80	6 items in Survey of Prices
0.57	0.67	0.32	0.45	0.12	CPI ex CE
0.63	0.75	0.32	0.59	0.14	CPI ex services & CE
1.54	-	-	0.64	0.90	Overall CPI vs overall FIES[7]

1990-2000

(C,C)[1]	(C,K)[2]	(K,C)[3]	(K,K)[4]	CPI-FIES[5]	
0.25	0.20	-0.54	-0.35	0.60	Food
0.72	0.73	-1.19	-1.39	2.09	Clothing
-6.12	-5.92	-2.89	-3.32	-2.80	Consumer Electronics
0.03	0.38	-0.81	-1.43	1.46	6 items in Survey of Prices
0.32	0.26	-0.46	-0.40	0.72	CPI ex CE
0.11	0.20	-0.81	-0.51	0.62	CPI ex services & CE
0.89	-	-	-0.46	1.35	Overall CPI vs overall FIES[7]

1 Original CPI fixed-weight Laspeyres index.
2 CPI price data and FIES monthly expenditure share used to compute a Tornqvist index.
3 CPI fixed weights and FIES unit prices used.
4 Tornqvist index using unit prices and expenditure shares from FIES.
5 Difference between CPI inflation rate and unit-price inflation rate in FIES.
6 1987-2000.
7 Baskets in two indices differ.

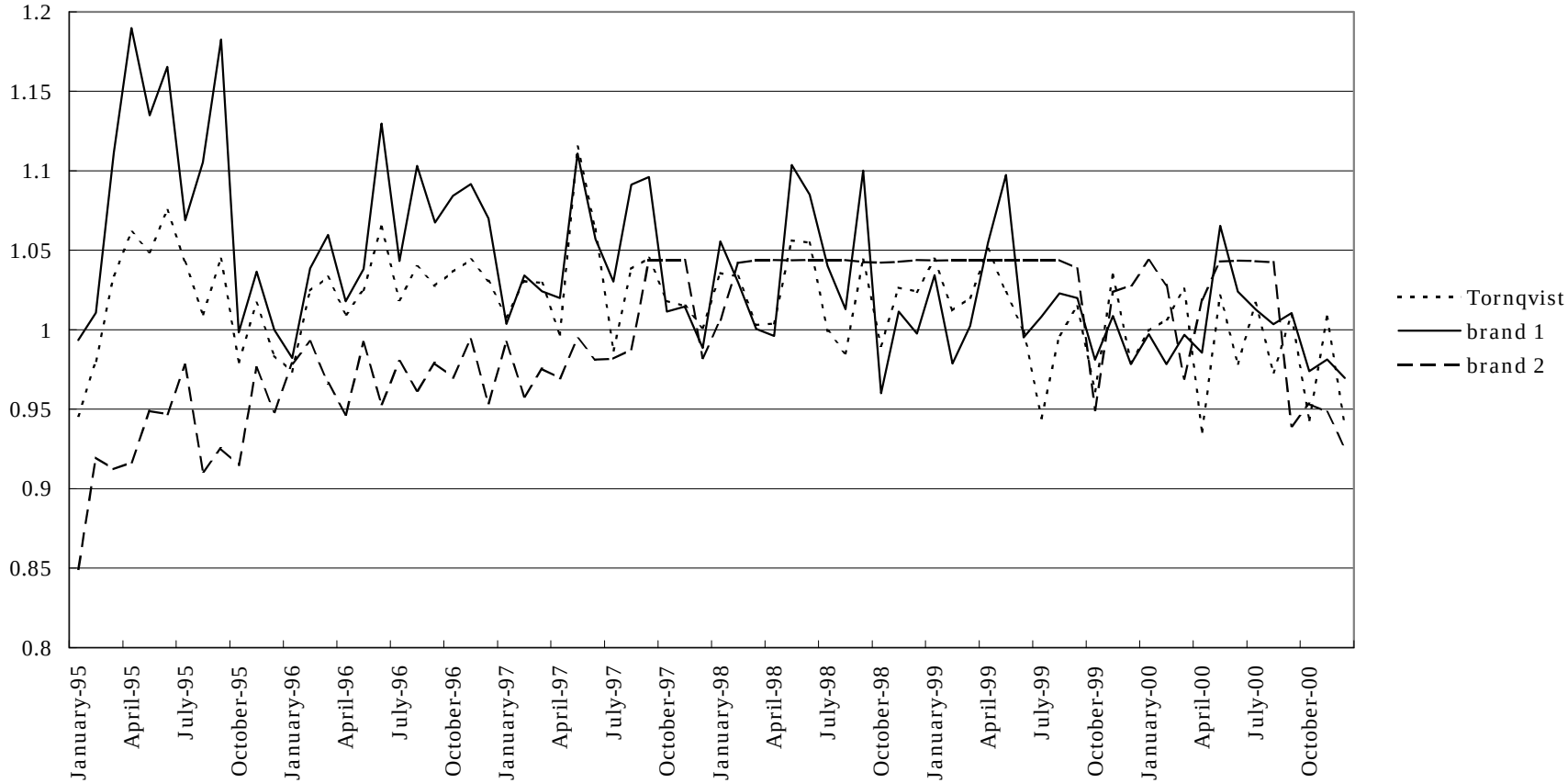
Table 13

Inflation Rates for 16 Selected Items

1995	1996	1997	1998	1999	2000	1995-2000 average	
-5.15	4.32	-7.66	0.72	-2.85	-6.47	-2.85	POS-Laspeyres
-5.05	3.17	-7.95	1.85	-2.80	-5.74	-2.75	POS-Tornqvist
-2.15	-2.07	-0.30	-1.00	-0.95	-2.28	-1.46	CPI

Notes The numbers shown are annual inflation rates (%). The first two use POS-SRI data (see Appendix 1 for the data source). The first row uses CPI weights and computes Laspeyres index, whereas the second is a chained Tornqvist using annual weights computed from the sales data in POS-SRI. The last row is computed using item level indices and respective weights in CPI.

Figure 1
Brand Specific Price Indices for Liquid Condiments



Appendix 1

Comments on the Data

Four sets of data are used in the analysis of potential mismeasurements of CPI.

1 Family Income and Expenditure Survey (FIES)

The CPI uses this survey for the selection and weights of items. Aside from the expenditure records of the roughly 8,000 sample households, FIES also reports average unit purchase prices for 200 items. We use these unit price data as the benchmark for cost of living index estimates.

There are three major problem with using these unit prices as the COLI. First, the data cover only subsets of consumption expenditure and do not cover services. Second, they are averages of nominal purchase prices without incorporating any changes in quality. Third, there are large monthly fluctuations in the data, partially reflecting measurement errors.

There are several advantages, however, over the current CPI as the benchmark of COLI. The unit price data reflect the average of the actual choices by sample households of items, brands, quality, and stores, thus incorporating substitutions by households across items, brands, quality, and different types of stores. To the extent quality changes not reflected in prices are not quantitatively important, the unit price and expenditure data provide the most natural measure of COLI. Another notable advantage is that the survey can be used to estimate COLI across different types of households: although the current CPI supplements include CPI series for several different types of households, they incorporate only the differences in weights across households (they use the common average prices taken from the Survey of Prices.)

2 1997 National Survey of Prices, Special Volume on Bargain Prices

This survey selected 16 items and collected cross-section data on regular and bargain (discount) prices across regions, types of stores, and variety of other attributes such as location

characteristics and store sizes). We use data for 14 of these 16 items. We deleted 2 items, eggs and beef, because of the potential large quality differences across samples.

3 POS-SRI

POS data compiled by the Sezon Research Institute (SRI) on 16 items for the 72 months January 1995 through December 2000 for 20 stores in Metropolitan Tokyo belonging to a national chain of supermarkets. The report provides monthly average prices and sales separately for regular price and discount sales. We use the same 14 items selected above.

4 POS-DEI

POS data are compiled by the Distribution Economics Institute of Japan includes 6 items among the 14 selected items above. The data are daily price and sales records for roughly 320 brand-store combinations for 24 months between April 1995 and March 1997.

Appendix 2

Shopping-Storage Model

Consider a household that consumes at constant rate c per day. Assume it visits a retailer each $1/s$ days. The price of the consumption good is randomly drawn from a known distribution $F(p)$. Normalize this price so the highest price is 1. Shopping costs are δ per visit, storage cost is ε per day per unit, and costs associated with stock-out are ignored. For simplicity, assume the same amount, q, is purchased on each visit if the price is below some threshold, level, $\hat{p}$. Since the amount purchased per visit must on average equal consumption (c) so

$$sqF(\hat{p}) = c \qquad \text{EQ. 8}$$

Thus the amount of purchase per visit is given by

$$q = \frac{c}{sF(\hat{p})} \qquad \text{EQ. 9}$$

The average time needed to consume the stored good is q/c. On average, the amount in storage is half the amount purchased, so the average storage cost per unit of time is

$$\frac{1}{2}\varepsilon q = \frac{c\varepsilon}{2sF(\hat{p})} \qquad \text{EQ. 10}$$

The household minimizes average (per unit of time) total cost by the choice of $\hat{p}$ and s, taking δ, ε, and F as given.

$$\underset{\{\hat{p},s\}}{Min}\cdot[c\int^{\hat{p}} pdF(p) + s\delta + \frac{c\varepsilon}{2sF(\hat{p})}] \qquad \text{EQ. 11}$$

The POS-DEI data set can be used to obtain an empirical price distribution for the simulation. The data include daily sales and price data for six items sold at 14 sample stores. Each item includes 20 to 30 different brands. The top five brands by unit sales are chosen from

each store for the simulation. The data span the two-year period 7 April 1995 to 7 April 1997. Daily price data are used to compute the kernel price density function for each brand, each item, and each store.[8]

The range of parameters we used in simulation are:

$c = .2$ (one unit of purchase is equal to 5 days' consumption)

$\varepsilon = .001 \sim .01$

$\delta = .05 \sim .14$

All are measured in rates per day. For example, $\varepsilon = .001$ is equivalent to depreciation at 0.1% per day if the good is purchased at the regular (high) price. Using the minimum wage in Okinawa (the lowest) of around ¥600 per hour to set the low end, and assuming about 1-2 hours for shopping, shopping cost per visit ranges from ¥500 to ¥1400, which translates to 5% to 14% of ¥10,000 of groceries. The upper limit correspond to roughly 2 hour minimum wage in Tokyo Metropolitan area (=708 yen per hour).

Table A2.1 shows that across-store variations in total shopping cost and average purchase price are large. For the top-five brands, total shopping cost varies by over 8% between store 8, the lowest, and store 2, the highest. For average purchase prices, the range is also more than 8% (between stores 8 and 2). If consumers (bargain hunters) choose to buy the cheapest, variations are even larger: more than 12% in total costs (stores 9 and 1), and close to 13% in average purchase prices (stores 4 and 1). Variations in total shopping costs are smaller than those for average purchase prices because volume shopping of discounted items increases inventory holding costs.

Notice that the two coops tends to be more expensive, especially for bargain hunters. This reflects the fact that periodic discounts are less common in those stores than in supermarket chains.

---

[8] The pricing patterns are not uniform and the optimal shopping behavior incorporating the periodic price discounts are highly complex. For simplicity we assume a random drawing of prices from the empirical price distribution. See Ariga, Matsui and Watanabe (2000) for the dynamics of pricing strategy and shopping behavior.

Across-store variations in pricing patterns alone can give rise to sizeable variations in shopping frequency and storage. The other side of this fact is that consumers with different shopping and storage costs choose different stores even if all the stores are identical except for the pricing policy. This follows from the large variation in optimal shopping and storage costs across stores even after controlling for unit shopping and storage costs. For example, when δ=0.05 and ε=0.001, the shopping cost for the top-selling brand of item 1 varies between 0.045 and 0.113 and storage cost varies between 0.023 and 0.057.

To demonstrate this, Table A2.2 shows the cost-minimizing choice of store as unit shopping and storage costs are varied for the top-selling brand of item 1. In this specific case, store 5 minimizes the total shopping cost for those with lower shopping and storage costs. For those with somewhat higher costs, store 2 becomes the best choice, reflecting the fact that the optimal shopping and storage policy for store 5 involves sizable purchase at occasional but deep discounts. At even higher shopping and storage costs, the optimal choice shifts to store 9.

This example is not exceptional. Among the 3000 simulation cases, each of the 14 stores is the cost-minimizing choice in at least one case, although store 10 has only one such case. Store 3 is the overall winner, being the best choice in 509 cases.

Table A2.1 Variations in Total Cost and Average Purchase Price across 14 Stores

	Total Cost		Average Purchase Price	
Sample Store	Top 5 Brands	Cheapest	Top5 Brands	Cheapest
1 Coop #1	+3.42%	+8.75%	+2.46%	+10.33%
2 Coop #2	+3.51%	+8.17%	+2.65%	+7.83%
3 National Chain A #1	-2.731%	-0.33%	-6.15%	+1.51%
4 National Chain A #2	-0.58%	-3.03%	-2.83%	-2.56%
5 Unknown	-2.02%	+6.62%	-6.55%	+8.16%
6 Unknown	+0.26%	-1.10%	-2.33%	-0.85%
7 Unknown	-0.71%	+8.06%	-2.66%	+9.66%
8 Unknown	-4.74%	+0.92%	-6.63%	+1.36%
9 National Chain B #1	-2.53%	-3.38%	-3.49%	-1.11%
10 National Chain B #2	+.00%	+2.68%	-1.68%	+6.64%
11 Regional Chain C #1	-4.02%	+2.44%	-6.60%	+6.67%
12 Regional Chain C #2	+0.19%	+2.23%	-1.10%	+4.87%
13 National Chain D #1	+0.23%	+3.00%	-0.61%	+4.20%

Note 1 Numbers shown are percentage differences from Store #14 (not shown) which is used as the benchmark.

Table A2.2 Optimal Store Choice for item #1, the top brand

		δ								
	.001	.002	.003	.004	.005	.006	.007	.008	.009	.010
.05	5	5	5	5	5	5	5	5	5	5
.06	5	5	5	5	5	5	5	5	5	5
.07	5	5	5	5	5	5	5	5	5	5
.08	5	5	5	5	5	5	5	5	5	2
ε .09	5	5	5	5	5	5	5	2	2	9
.10	5	5	5	5	5	5	2	9	9	9
.11	5	5	5	5	5	2	9	9	9	9
.12	5	5	5	5	2	9	9	9	9	9
.13	5	5	5	2	9	9	9	9	9	9
.14	5	2	9	9	9	9	9	9	9	9

Note 1 The optimal choice of store under each configuration of ε and δ is shown in each cell.