Was Inflation Observed under the First Wave of the COVID-19 Spread in Japan? Scanner Data Evidence for Retailers in Tokyo

By

Masahiro Higo (The University of Tokyo)
Shigenori Shiratsuka (Keio University)

September 2022

CREPE DISCUSSION PAPER NO. 131

CRePE
Center for Research and Education in Program Evaluation

CENTER FOR RESEARCH AND EDUCATION FOR POLICY EVALUATION (CREPE)
THE UNIVERSITY OF TOKYO
http://www.crepe.e.u-tokyo.ac.jp/

Was Inflation Observed under the First Wave of the COVID-19 Spread in Japan? Scanner Data Evidence for Retailers in Tokyo [*]

Masahiro Higo[†] and Shigenori Shiratsuka[‡]

September 6, 2022

Abstract

In this paper, we examine whether inflation was observed under the first wave of the COVID-19 spread. To address this issue, we construct high-frequency quality-adjusted price indices by employing daily scanner data of retail stores in Tokyo. We attempt to make explicit adjustments for not only the product characteristics but also the structural changes in temporary sales and the retail service quality of outlet channels. We emphasize that adjustments for the effects of temporary sales and retail service quality are particularly important in examining the retail price dynamics under the COVID-19 pandemic as the voluntary lockdown constrained household purchasing behavior because of the risk of COVID-19 infection. We conclude that mild and temporary inflation of slightly less than 1% was observed during the first wave of the COVID-19 spread. Based on the estimation results, we decompose the differences between the increases in the unit prices and those in the quality-adjusted price indices into outlet substitution effects and temporary sales effects.

Keywords: Consumer price index, Scanner data, Quality adjustment, Retail service quality, Temporary sales, COVID-19

JEL codes: C43, E31,

[*] This paper is prepared for the special issue for Japan and World Economy on "Government Statistics and Economic Performance." INTAGE Inc. provided the daily scanner data used in this paper under the collaborative research project at the Center for Research and Education in Program Evaluation (CREPE), the University of Tokyo. We thank Erwin Diewert, Shin'ichi Fukuda, Taisuke Nakata, Tatsuyoshi Okimoto, Mototsugu Shintani, Nao Sudo, Jiro Yoshida, and participants of the Japan and World Economy Symposium and seminar at the Bank of Japan as well as an anonymous referee for their constructive comments and discussions. Shiratsuka acknowledges the financial support from Keio University Academic Development Fund.

[†] The University of Tokyo. E-mail: masahiro.higo@e.u-tokyo.ac.jp

[‡] Keio University. E-mail: shigenori.shiratsuka@keio.jp

1 Introduction

In this paper, we examine whether inflation was observed under the first wave of the COVID-19 spread. To address the question, we construct high-frequency quality-adjusted price indices by employing daily scanner data at retail stores in Tokyo.

During the first wave of the COVID-19 spread in Japan in the first half of 2020, the number of infections began to increase from mid-February and spiked through March. In response, the Government of Japan declared a state of emergency on April 7 in seven prefectures, including Tokyo, and on April 16, the state of emergency was extended to the entire country. Under these circumstances, people refrained from going out. The number of new infections peaked in mid-April and began to decline rapidly. During the first wave of the COVID-19 spread, Japan experienced weaker social and economic constraints under the state of emergency. Unlike the strict lockdown in many countries, Japan implemented a voluntary lockdown, experiencing a significant decline in economic activities.

Many previous studies have used various types of alternative data in Japan to discuss the social and economic issues during the COVID-19 pandemic.[1] For example, Hosono (2021) explored the low infectious cases and the large decline in consumption by extending the SIR-macro-model to incorporate Japan's two key factors of voluntary stay-at-home and a request-based lockdown. Using smartphone location data, Watanabe and Yabu (2021) revealed that both the intervention effect of government policy and the information effect of health risks from COVID-19 infection constrain people's behavior in response to the stay-at-home measures. Shibamoto et al. (2022) estimated a dynamic model of the interaction between infection-mobility tradeoff and mobility demand. Fujii and Nakata (2021) pointed out that the tradeoff between output and infection in the short run does not necessarily exist in the long run based on their simulation results with the SIR-macro-model with time-varying parameters.

Among such studies, some have employed various types of alternative data to analyze the structural changes in household consumption expenditure and retail sales trends. These studies did not only use scanner data for retail stores but also credit card expenditure history and data from household accounting applications. For example, Watanabe and Omori (2020) and Konishi et al. (2021) revealed that household consumption expenditure patterns changed significantly, especially from eating out to home cooking and delivery services. As a result, the retail sales trends also changed drastically, observing that supermarkets enjoyed positive impacts, whereas department stores and restaurant affiliates were negatively affected. However, only a very limited number of studies, including the mask price analysis in Abe et al. (2020), focused on retail price development under the COVID-19 pandemic.[2]

[1] Japanese Economic Association has set up a webpage for information on academic papers analyzing the impact of COVID-19 (https://covid19.jeaweb.org/scientific.html).

[2] Retail scanner data research in Japan dates back to the mid-1990s when the Seiyu price index, compiled by one of the major supermarket chains in Japan, highlighted the "price busting" phenomena in Japanese retail markets. In the mid-1990s, Japan experienced a relatively high domestic price level. With continued low inflation under the "lost decades," Japan's price problem turned into a relatively low domestic price level. Scanner data research focuses on stealth inflation with product turnover and downsizing (Imai and Watanabe, 2014; Ueda et al., 2019; Ueda

Focusing more on the measurement issues of the consumer price index (CPI) under the COVID-19 pandemic, Diewert and Fox (2020) pointed out the possible downward bias, not upward bias, in the CPI due to the sudden unavailability of goods and services during the pandemic. Cavallo (2020) attempted to quantify the downward bias in the CPI due to changes in consumer expenditure patterns by employing credit and debit transaction data. These studies mainly focused on the measurement bias of the CPI stemming from the aggregation formula with the fixed consumption basket.

The COVID-19 pandemic may lead to mismeasurement of individual prices due to inadequate adjustments for the temporary sales effects and retail service quality. The same product is sold at different prices at different outlets as observed prices reflect the differences in the temporary sales effects and the retail service quality across outlets. For example, prices at convenience stores are generally higher than those at other outlet channels, reflecting their convenience in terms of operating hours, location, and shopping time. The sudden and unexpected changes in consumer expenditure patterns induced by the COVID-19 spread are likely to produce the measurement bias in the CPI caused by individual prices without making an explicit adjustment for the temporary sales effects and the differences in retail service quality.

Fig. 1 depicts the recent development of retail prices using a high-frequency scanner price index for food and beverage products and daily necessities at supermarkets.[3] The T-index indicates a moderate upward trend from shortly before the first declaration of a state of emergency in April 2020. The trend then reversed after reaching a peak under the state of emergency, revealing a downward trend toward mid-2021. During this period, two observations are noted. First, the volatility of the T-index declined significantly, suggesting possible declines in the frequency of temporary sales and the size of price reduction. Second, the T-mode-index deviated from the upward trend of the T-index significantly, although the T-index and the T-mode-index experienced similar movements outside this period, indicating that increases in the T-index are attributed to decreases in the frequency and the size of price reductions of temporary sales.[4] Moreover, the voluntary lockdown under the state of emergency induced people to minimize shopping frequency, travel distance, and shopping time, influencing household purchasing behavior, such as the choice of retail outlets and price sensitivity of consumption expenditure decisions. Fig. 2 depicts no significant changes in household perception of inflation during the first half of 2020, including the first wave of the COVID-19 spread.

et al., 2019), which tries to raise unit prices by reducing the product volume while keeping tag price unchanged. Moreover, applications of scanner data have continued to expand by linking more broad aspects of macroeconomic issues, such as the frequency of price changes to explore the flattened Phillips curve (Abe and Tonogi, 2010) and the interaction of the frequency of temporary sales with macroeconomic conditions and effectiveness of monetary policy (Sudo et al., 2018).

[3] NIKKEI CPINow is the real-time daily price index computed from scanner data of supermarkets. It covers food and beverage products and other daily necessities with a fairly long time series from 1989.

[4] The T-mode-index is computed from mode prices for individual products at specific stores with daily windows from 28 days before to 28 days after.

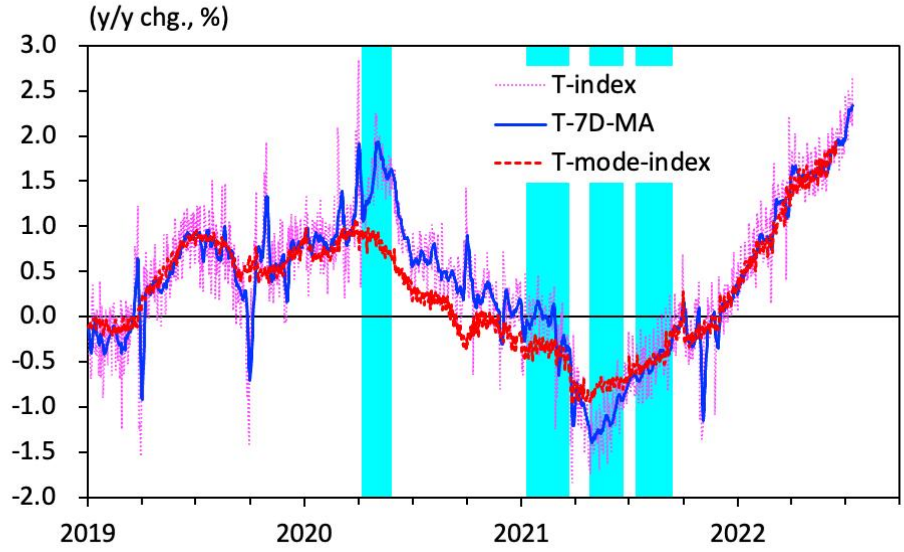


Fig. 1. Daily Retail Price Developments

Notes: The T-index aggregates product prices for each outlet at a JAN code level by applying the Törnqvist index formula. The T-7D-MA is a seven-day backward moving average of the T-index. The T-mode-index employs mode prices for the daily windows from 28 days before to 28 days after. The light blue areas indicate the periods under the declaration of states of emergency in Tokyo—(i) from April 7 to May 25, 2020, (ii) from January 8 to March 21, 2021, (iii) from April 25 to June 20, 2021, and (iv) from July 12 to September 30, 2021. Year-on-year changes are calculated as percentage changes from 364 days earlier to adjust the week-day effects. The plotted data for CPINow are nationwide, as it does not publish the data for Tokyo metropolitan area, though it publishes other prefectural data.

Source: Nowcast, *NIKKEI CPINow*.

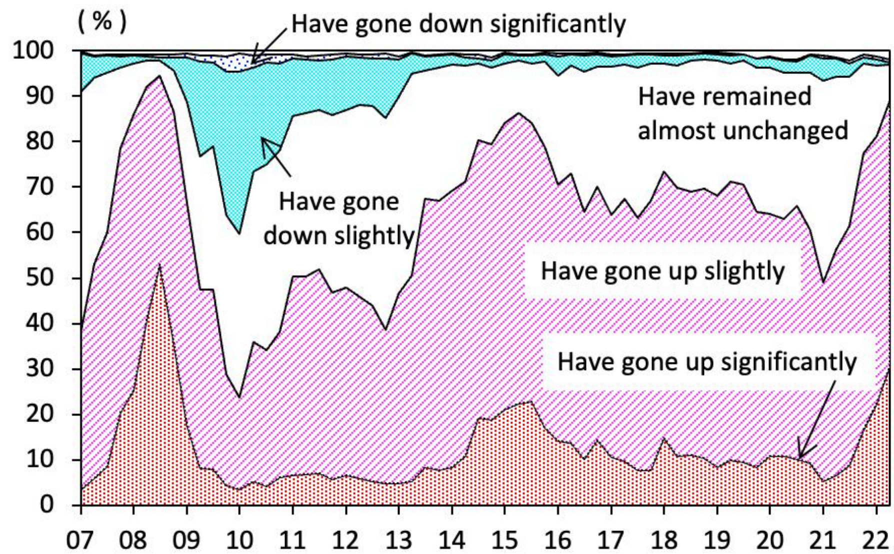


Fig. 2. Inflation Perception by Household

Source: Bank of Japan, *Opinion Survey on the General Public's Views and Behavior*.

Fig. 3 depicts that CPI inflation for the headline indicator and three core indicators, that is, headline less perishables, headline less perishables and energy, and trimmed mean, experienced a downward trend in the first half of 2020. However, the Japanese CPI fails to reflect the structural changes in the retail markets, probably due to its price survey method based on the "one-specification-for-one-item" policy. The policy specifies a few popular specifications for each item and continuously surveys their prices at specific outlets. Therefore, it is difficult to maintain price representativeness under the rapid and significant structural changes in the retail markets.[5]

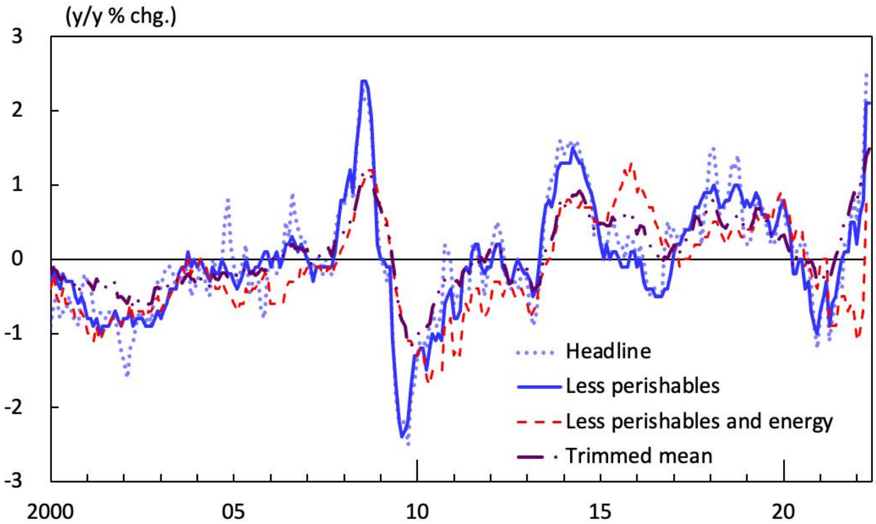


Fig. 3. Consumer Price Inflation in Japan
Sources: Statistics Bureau of Japan, *Consumer Price Index*, Bank of Japan, *Measures of Underlying Inflation*.

The following question arises: “Are the unit price increases observed in Fig. 1 considered price changes on a quality-adjusted basis in terms of not only product characteristics but also retail services at outlets, including the temporary sales effects?” We construct high-frequency quality-adjusted price indices to address the question by making explicit adjustments to daily scanner data on product characteristics, temporary sales effects, and retail service quality. As the voluntary lockdown constrained household purchasing behavior, we emphasize that quality adjustments for retail services, including the temporary sales effects, are particularly important in examining the retail price dynamics during the COVID-19 pandemic. Households responded to the voluntary lockdown by reducing shopping frequency, travel distance, and shopping time. Therefore, a quality-adjusted price index is considered to exist between the CPINow T-index, which fully reflects the decrease in temporary sales as a price increase, and the T-mode-index, which excludes the temporary sales effects.

Our approach is closely related to de Haan and Hendriks (2013), who applied the time-product dummy (TPD) method to daily web-scraping data of Dutch online stores.[6] They com-

[5] As Shiratsuka (2021) pointed out, the Japanese CPI has problems with the broadly-defined lower-level substitution bias, which corresponds to the weakened price representativeness stemming from the "one-specification-for-one-item" policy of the price survey method.

[6] de Haan and Hendriks (2013), de Haan and Krsinich (2014), and de Haan (2015) proposed two estimation procedures to incorporate scanner and web-scraping data into price indices—the time dummy hedonic (TDH) index (when information on item characteristics is available) and the TPD index (when unavailable). Krsinich (2016) examined the application of the TPD index combined with rolling estimations, called the fixed effects window splicing (FEWS) index. She argued that the FEWS is robust to retroactive revisions from data accumulation over time, but it is less applicable to products with rapid changes in product characteristics and consumer preferences. Haan et al. (2021) pointed out that the TPD index is distorted toward zero due to overfitting. They argued that

pared the TPD index with other conventional price index formulas, indicating downward biases with volatile price trends. We extend the approach of de Haan and Hendriks (2013) in three respects. First, we extend the TPD model by taking explicit account of the difference in the retail service quality of outlet channels (the extended TPD model). We employ the combinations of the two fixed effects, i.e. products and outlet channels, in the cross-sectional direction as our dataset includes sales records of various store types in the same product. Second, we construct a high-frequency quality-adjusted price index on a weekly basis using daily scanner data. Due to the strong weekly effects observed in the retail sales trends, we construct weekly quality-adjusted price indices based on the estimates for weekly dummies applied to both daily and weekly converted data. Third, we apply the extended TPD model to all food and beverage products in a unified manner, thereby aggregating them into the overall food and beverage product price indices.

Our conclusions are summarized as follows. We reveal that mild and temporary inflation of slightly less than 1% (0.6−0.9%) was observed during the first wave of the COVID-19 spread. We also decompose the differences between the increases in the unit prices and those in the quality-adjusted price indices into outlet substitution effects and temporary sales effects. In the daily frequency estimation, the unit price increase of 1.7% is decomposed into temporary sales effects of 0.3%, outlet substitution effects of 0.5%, and quality-adjusted price increase of 0.9%, whereas in the weekly frequency estimation, it is 0.4%, 0.7%, and 0.6%, respectively.

The rest of this paper is as follows. Section 2 describes the extension of the TPD model to incorporate the explicit quality adjustment for retail service quality. Section 3 explains the INTAGE SRI (Nationwide Retail Store Panel Survey) scanner data and summarizes the retail sales trends around the first wave of the COVID-19 spread. Section 4 explains our quality adjustment strategy and constructs high-frequency quality-adjusted price indices based on the baseline estimation results for the extended TPD model. Section 5 describes comprehensive robustness checks on the baseline estimation results. Section 6 discusses whether inflation was observed under the first wave of the COVID-19 spread in Tokyo. Finally, Section 7 concludes the paper by discussing the implications of our estimation exercises for the official CPI compilation methods.

2 Empirical Framework

We extend the TPD model in the studies by de Haan and Hendriks (2013) and de Haan and Krsinich (2014) by explicitly considering the difference in the retail service quality of outlet channels (extended TPD model). Assuming that N different products are sold during the sample period from 0 to T, the period 0 is the base period, and the periods from 1 to T are the comparison periods. $p_{i,t}$ denotes the price of a product i at period t ($t = 0, 1, ..., T$). By using log-transformed prices as a dependent variable, the log-linear time-dummy hedonic (TDH)

the TDH method makes robust and explicit quality adjustments based on quality characteristics, whereas the TPD method has problems with making implicit quality adjustments.

model is stated as follows:

$$lnp_{i,t} = \alpha + \sum_{t=1}^{T} \lambda_t TD_{i,t} + \sum_{k=1}^{K} \beta_k z_{k,i} + \epsilon_{i,t}, \tag{1}$$

where $TD_{i,t}$ is the time dummy variable, which takes the value of 1 if the observation is the period t and 0 otherwise. $z_{k,i}$ is the indicators of product characteristics k for a product i, and $\beta_k$is the corresponding estimated parameter. $\epsilon_{i,t}$ is the error term. The price index from period 0 to t is computed by taking the exponential of the estimated parameter for the time dummy λ_t.

Similarly, the log-linear TPD model is defined as follows:

$$lnp_{i,t} = \alpha + \sum_{t=1}^{T} \lambda_t TD_{i,t} + \sum_{i=1}^{N-1} \gamma_i PD_i + \epsilon_{i,t}, \tag{2}$$

where PD_i is the product dummy variable, which takes the value of 1 if the observation relates to a product i and 0 otherwise, and the dummy for an arbitrary product N is excluded in the estimation specification to identify the model. The estimated parameters λ_t andγ_i are fixed effects for time and product, respectively. As with the TPD model, the price index from period 0 to t is computed by taking the exponential of the estimated parameter for the time dummy λ_t.

We extend the log-linear TPD model by making an explicit quality adjustment for retail services by introducing an additional cross-sectional control variable for stores as follows:

$$lnp_{i,t} = \alpha + \sum_{t=1}^{T} \lambda_t TD_{i,t} + \sum_{i=1}^{N-1} \gamma_i PD_i + \sum_{s=1}^{S-1} \delta_s ST_{s,i} + \epsilon_{i,t}, \tag{3}$$

where $ST_{s,i}$ is the store dummy variable, which takes the value of 1 if a product i is sold at a store s and 0 otherwise, and the dummy for an arbitrary store S is excluded in the estimation specification to identify the model.

As discussed by Krsinich (2016), in theory, the price indices based on the TDH model and the TPD model are the same because the fixed effect for a product i corresponds to the net effect of the characteristics for a product i. However, in practice, the TPD model is less efficient than the TDH model if sufficient information on the characteristics is available. The number of relevant characteristics can be narrowed down considerably by applying the TDH model.[7] However, the TDH model is often difficult to apply to analyses with scanner data because, in many cases, it is hard to link scanner data to a dataset of the characteristics of a wide range of products.

Thus, choosing either the TDH or TPD model is a practical issue based on data availability. We consider the TPD model a realistic option in this paper because it needs to handle scanner data covering a wide variety of product categories in a unified manner.

[7] For example, see the study by Shiratsuka (1995) for the selection of explanatory variables in the THD model. He applied the TDH model to automobiles in Japan by selecting three key performance characteristics—max power, wheelbase, and interior space—out of 11 characteristics.

3 Data

In this section, we explain the INTAGE SRI scanner data and then summarize the retail sales trends under the first wave of the COVID-19 spread.

3.1 INTAGE scanner data

We employ daily scanner data of food and beverage products, SRI (Nationwide Retail Store Panel Survey) provided by INTAGE Inc. The dataset contains the daily sales, size, and quantity information of nationwide retail stores in Japan at a barcode level from January 2019 to June 2020. The sales records cover sales of food and beverage products with JAN codes, recorded as a stock keeping unit (SKU) in the dataset. Food and beverage products cover about 20% of the consumption expenditure basket for the CPI. The outlets are classified into seven types−(i) general merchandise stores (GMS), having a sales floor of 3,000 or more square meters and more than 50 employees, (ii) supermarkets (SM), having 500 square meters or above, (iii) mini supermarkets (M-SM), (iv) convenience stores (CVS), (v) home centers and discount stores (HC/DS), (vi) large drug stores (Drug/L), and (vii) discount liquor stores (Liquor/DS).[8]

We focus on the data for retail stores in Tokyo from January to June 2020 to analyze the effects of the first wave of the COVID-19 spread in Tokyo. Table 1 summarizes information on the dataset from the product side on the upper panel and the store side on the lower panel. The upper panel reveals that the total number of SKUs is just over 76,000, categorized into 151 items, such as rice and bread. The items are further classified into eight categories, including staple foods. The lower panel presents the daily average data on retail trends. GMS has the largest number of price observations (products) and attracts many customers, registering large sales amounts in a relatively small number of stores. SM has the largest total sales amounts and a relatively large number of price observations and customers. In contrast, CVS and Drug/L have a limited number of price observations and relatively low total sales amounts in a relatively large number of stores.

[8] INTAGE data divide SM into large and small sizes by floor size, but we merge them as similar retail sales trends are observed in the two outlet types. We also drop outlets classified as small drug stores (Drug/S) and small liquor shops (Liquor/S) because of the small outlet samples and fewer SKUs. INTAGE scanner data expand its coverage of stores from January to March 2019, especially in HC/DS.

Table 1
Summary Statistics

(1) Items and SKUs

Category	Number of items	Number of SKUs	SKUs per item		
			Mean	Maximum	Minimum
Staple foods	17	10,580	622	3,204	30
Seasonings	36	11,890	330	1,696	11
Processed foods	45	20,220	449	2,117	22
Sweets / Snacks	19	16,270	856	2,695	10
Milk beverages	4	789	197	441	55
Coffee / Tea	8	2,049	256	706	61
Soft drinks	16	5,074	317	724	73
Alcoholic beverages	6	9,058	1,510	2,744	491
Total	151	75,930	503	3,204	22

(2) Retail Sales Trend (Daily Average)

Outlet type	Sales amount (million yen)	Number of stores	Per outlet		Per customer
			Number of price observations	Number of customers	Sales amount (yen)
GMS	71.8	15.1	4,641.0	6,505.7	743.8
SM	126.0	60.9	2,927.8	2,639.8	783.9
M-SM	13.6	17.6	1,469.5	1,297.8	586.3
CVS	7.3	51.4	473.7	728.1	204.1
HC/DS	12.9	14.6	909.7	2,510.3	335.9
Drug/L	12.9	85.6	380.3	704.0	193.6
Liquor/DS	4.7	9.7	406.5	335.2	1,413.0
Total	249.2	254.9	1,366.4	1,645.5	450.7

Notes: Figures in the table include SKUs sold from January 6, 2020, to June 28, 2020.

3.2 Unit price and effective quantity indices

We employ unit prices as the dependent variables in estimating high-frequency quality-adjusted price indices at an elementary aggregation level. Prior to our analysis, we examine how the relationship between unit prices and quantity changed across retail channels during the first wave of the COVID-19 spread.

The sales amounts are calculated by multiplying the price (P) and the quantity (Q) of an individual product and summing them up over all products. The price is decomposed into the product of the unit price (UP) and unit size (US). The unit size multiplied by the quantity is defined as the effective quantity (EQ). Thus, the sales amounts can be redefined as the summation of the product of the unit price and effective quantity of all products, as follows:

$$SA_{k,t} = \sum_i P_{i,k,t} \times Q_{i,k,t} = \sum_i UP_{i,k,t} \times US_{i,k,t} \times Q_{i,k,t} = \sum_i UP_{i,k,t} \times EQ_{i,k,t}, \quad (4)$$

where subscript i, k, and t indicate a product, store, and time, respectively.

Based on the decomposition of Eq. (4), we construct the unit price index (UPI) and effective quantity index (EQI) using the real-time weight of the daily sales amounts as shown by Eqs. (5) and (6) below:

$$UPI_t = \sum_i \sum_k sw_{i,k,t} UPI_{i,k,t}, \tag{5}$$

$$EQI_t = \sum_i \sum_k sw_{i,k,t} EQI_{i,k,t}, \tag{6}$$

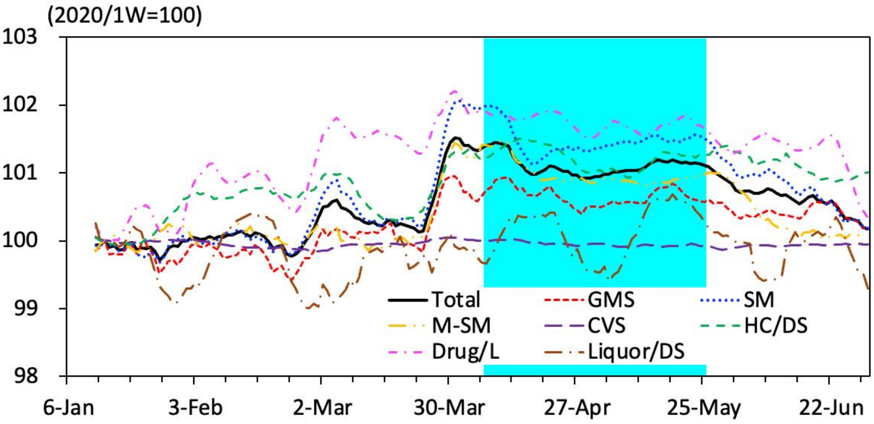
where sw denotes weight for a product i at store k at time t. The indices are constructed using the first week of January 2020 as the base period.

Fig. 4 depicts the computed results for the UPI and EQI across store types. Three points should be noted. First, the sharp increase in the UPIs for GMS, SM, M-SM, and HC/DS shortly before the first declaration of the state of emergency corresponds to the spike in the EQIs. The UPI for Drug/L increased slightly in early March. The UPIs for other store types did not experience such significant increases, especially those for CVS experienced a very flat trend over time, even under the first declaration of the state of emergency. Second, the EQIs for most store types, except for CVS and Liquor/DS, experienced two spikes in the early and end of March, suggesting a stock-piling behavior at many households in preparation for the stay-at-home request under the state of emergency. Third, the UPIs for GMS, SM, and M-SM, which increased shortly before the first declaration of the state of emergency, started declining after lifting the state of emergency.

Based on the first and second points, the increases in the UPIs are associated with the increases in the EQIs, suggesting that all the increases in the UPIs should not be considered as price hikes on a quality-adjusted basis. During the first wave of the COVID-19 spread, due to the increased risk for COVID-19 infection, households decided to purchase at higher prices than usual to reduce shopping frequency, travel distance to stores, and shopping time.[9]

[9] As can be seen from Fig. 4, the effective sales value remained flat at a high level during the first declaration of the state of emergency. Although not depicted in the figure, the number of customers visiting stores decreased during this period, whereas the sales amounts per customer increased. This observation also suggests that households tried to minimize shopping frequency by making bulk purchases in a single shopping trip to control the risk of COVID-19 infection.

(1) Unit Price Indices



(2) Effective Quantity Indices

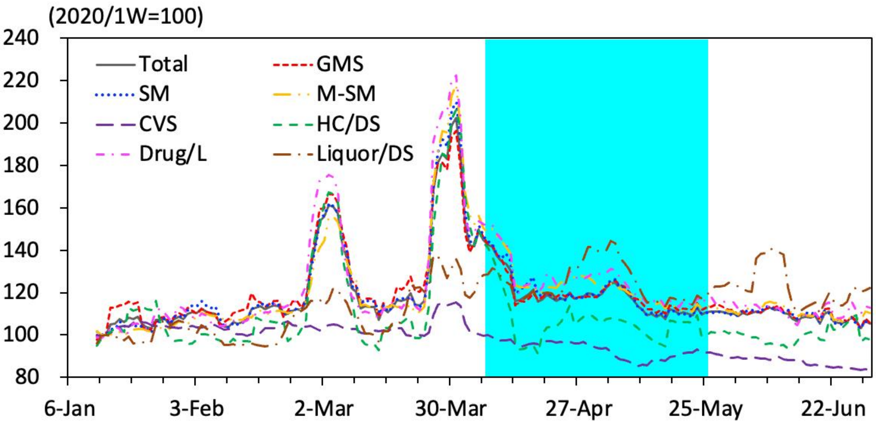


Fig. 4. Unit Price Indices and Effective Quantity Indices by Outlet Types
Notes: The plotted figures are backward seven-day moving averages, starting from the date on the horizontal axis. The light blue areas indicate the period of the first declaration of a state of emergency in Tokyo.

3.3 Temporary sales

We now examine the effects of temporary sales. We define temporary sales as a day that the observed price of a product at a store is at least two yen lower than the mode price for the past seven days, including the current day, following the study of Abe and Tonogi (2010). We also define a price reduction as the deviation of the observed price at a temporary sale from the mode price. We aggregate the frequency and size of price reductions over items by computing simple mean and weighted mean with sales amounts.

Fig. 5 depicts the frequency of temporary sales and the size of price reductions at temporary sales over time.[10] These figures reveal that temporary sales declined rapidly from early March

[10] Sudo et al. (2018) examine the effects of temporary sales on the macroeconomic implication by focusing on the households' allocation of time for work, leisure, and bargain hunting. They compute various measures for the frequency of temporary sales based on two types of sales filters: mode price during a certain window length (Eichenbaum et al., 2011, and Kehoe and Midrigan, 2015) and filtered V-shaped price fluctuations as a temporary sale (Nakamura and Steinsson, 2008). Their estimates for the closest definition to ours, mode price for a seven-day widow, are around 0.2 in the 2010s and slightly higher than our estimates in early 2020. In this respect, our estimates for the frequency of temporary sales become slightly higher, around 0.2, when limiting data only for GMS and SM, which seems consistent with the estimates in Sudo et al. (2018).

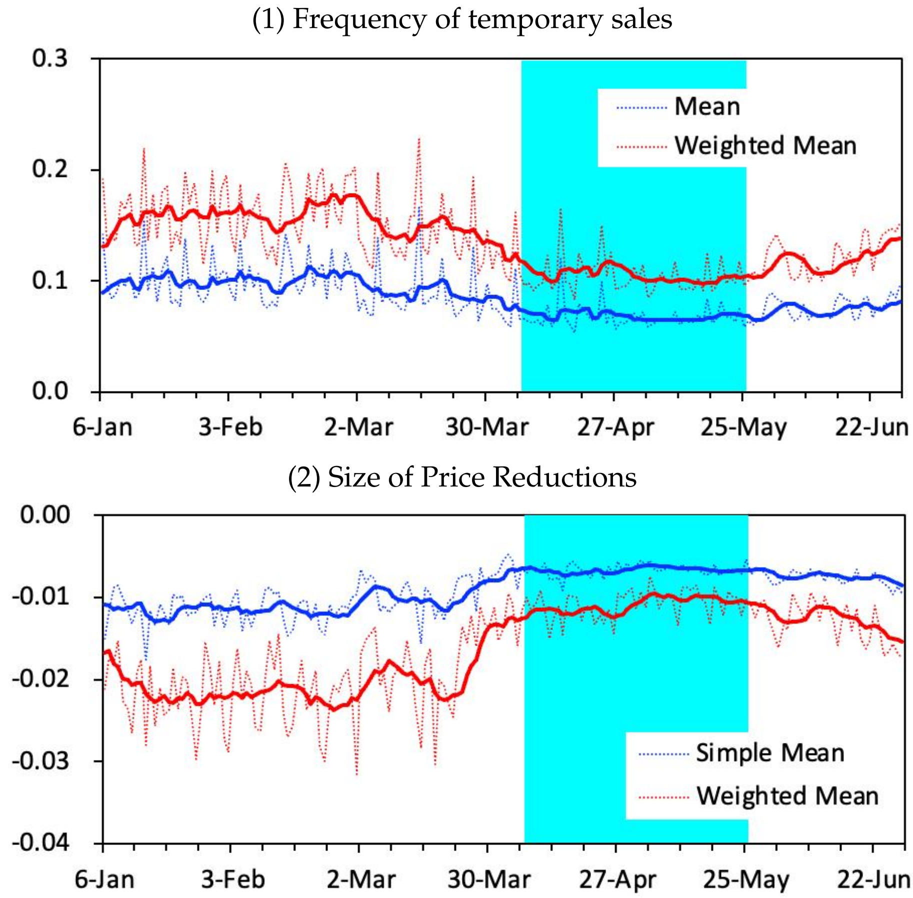


Fig. 5. Frequency and Price Discount of Temporary Sales
Notes: The bold lines are backward seven-day moving averages starting from the date on the horizontal axis. The light blue areas indicate the period of the first declaration of a state of emergency in Tokyo.

toward the first declaration of the state of emergency, as indicated by the significant divergence between the T-index and the T-mode-index of CPINow in Fig. 1. Before March, the differences between the simple and weighted means remained almost constant, but they started narrowing in early March. These observations confirm that both the frequency of temporary sales and the size of price reductions declined during this period. This suggests the importance of considering the temporary sales effects on the retail sales trends when computing the high-frequency quality-adjusted price indices.

To confirm the structural changes in temporary sales under the first wave of the COVID-19 spread, we perform a statistical test for structural changes with an unknown break point. Fig. 6 plots the sup-Wald statistics for structural changes in the differences between simple and weighted means for the share of temporary sales and the size of price reductions, respectively. The sup-Wald statistics are computed from January 6 to June 28 in 2020, with 15% trimming for detecting an unknown break point. The sup-Wald statistics start increasing from mid-February and reach their peaks around the end of March. The detected breakpoints for the share of temporary sales and the size of price reductions are March 28 and 24, 2020, respectively.

It should be noted that declines in the frequency of temporary sales and the size of price reduction influence how to measure the central tendency of the unit price distributions over time.

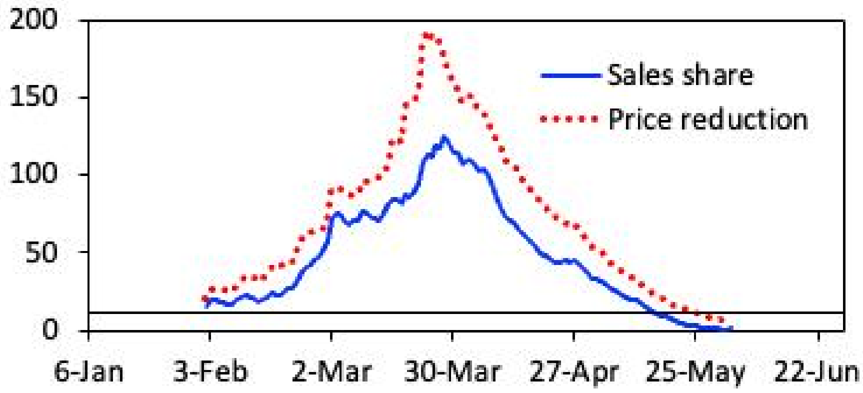


Fig. 6. Test for Structural Changes with an Unknown Breakpoint
Notes: The plotted figures are the sup-Wald statistics for detecting an unknown break point for the differences between the simple and weighted means for sales share and price reductions. The bold horizontal line indicates the critical value for 1% with 15% trimming for detecting an unknown break point. The breakpoints for sales share and price reductions are March 28 and 24, 2020, respectively.

Fig. 7 plots three statistical measures for the central tendency of the unit price distributions over time—mean, median, and mode of weekly windows for a product at a store—aggregated by the time-varying sales amount weights. The three measures are indexed at 100 with the mode values for the first week of January 2020.

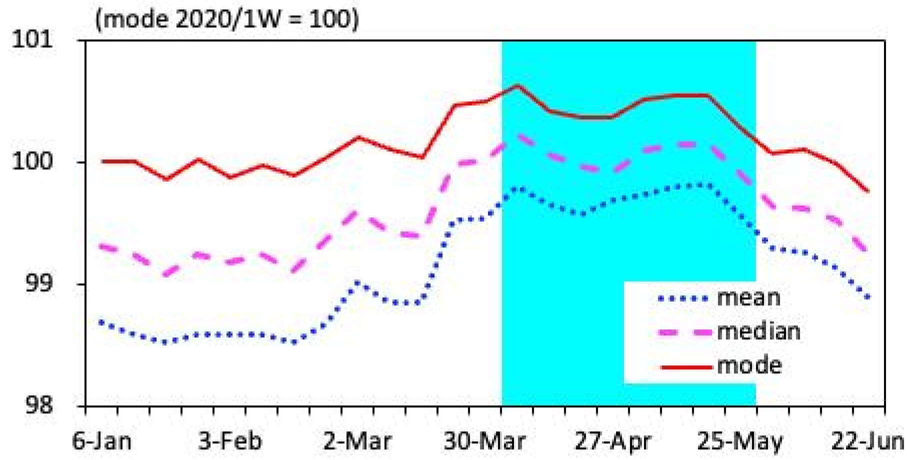


Fig. 7. Summary Statistics for Central Tendency of Unit Price Distributions
Notes: The light blue area indicates the period for the first declaration of a state of emergency in Tokyo.

The order of the three statistics remains unchanged over time: mode, median, and mean from top to bottom, indicating that the unit price distributions are skewed toward the left. However, the mean moves closer to the mode, reflecting a decrease in the temporary sales effects from the end of February, and the leftward-skewness of the unit price distributions also declined. This implies that inflation trends are estimated differently, depending on the measures for the central tendency of the unit price distributions used. The mean values for unit prices are most influenced by the declines in temporary sales under the first wave of the COVID-19 spread, and the mode values are the least influenced.

However, the unit size of products purchased by households did not change significantly during the first wave of the COVID-19 spread. Fig. 8 plots quantiles of the unit size distributions for products purchased by households on a weekly basis. The plotted figures are normalized using the mean and standard deviation for the first week of January 2020. The distributions remain almost unchanged, slightly shifting toward a larger direction. In response to the declined temporal sales, households continued to purchase products with approximately the same unit size, suggesting that, due to the increased risk for COVID-19 infection, households chose to purchase at higher prices to reduce travel distance to stores and shopping time.

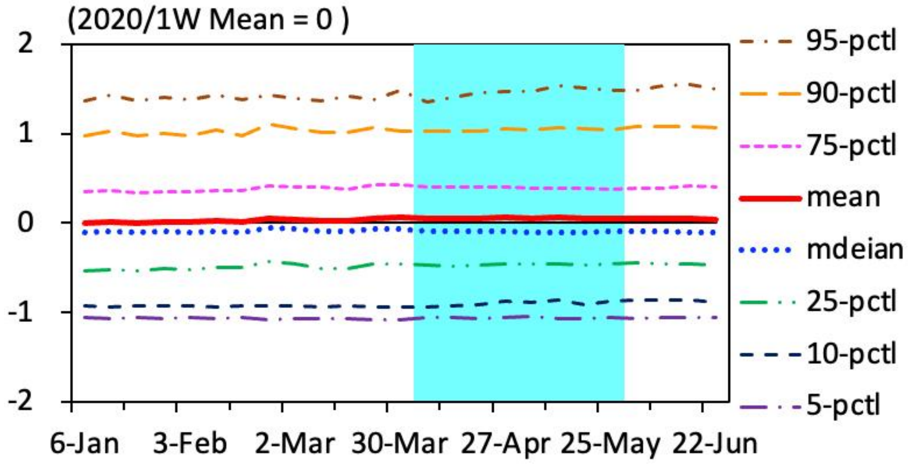


Fig. 8. Unit Size Distributions
Notes: The light blue area indicates the period of the first declaration of a state of emergency in Tokyo.

4 High-frequency Quality-adjusted Price Index

In this section, we explain our quality adjustment strategy and construct high-frequency quality-adjusted price indices based on the baseline estimation results for the extended TPD model.

4.1 Quality adjustment strategy

Given the observations about the significant structural changes in the retail sales trends under the first wave of COVID-19 in the previous section, we point out the importance of the quality adjustments in two respects, in addition to the standard quality adjustments for product characteristics. First, we need to pay careful attention to accounting for the structural changes in the frequency and size of price reductions of temporary sales. Due to the significant structural changes in the relative relationship between the regular and discount prices of temporary sales, the treatment of discount prices significantly influences the estimated price index. It should be noted that both supply- and demand-side factors induce such structural changes in the retail sales trends. On the one hand, retailers try to raise prices in real terms by reducing the frequency and size of price reductions of temporary sales to cover the increased costs of infection prevention measures. On the other hand, consumers are more likely to purchase at higher-than-normal prices to minimize the risk of COVID-19 infection under voluntary lockdown.

Second, we also need to consider outlet substitution bias. This bias is induced by the structural change in retail markets, as observed by the shift in shopping sites across the outlet types. The shifts in nearby outlets indicate possible changes in consumers' preference due to the stay-at-home request under the voluntary lockdown and the risk of COVID-19 infection. Under such circumstances, households are likely to alter their shopping behavior to minimize shopping frequency, travel distance, and shopping time. Such changes in consumers' behavior indicate that, on a quality-adjusted basis, they perceive that the prices for the same products sold at supermarkets nearby are less expensive than before.[11]

To deal with the temporary sales effects and the retail service quality differences, we employ four baseline specifications in Table 2—D-ROL_SK+ST, D-ROL_SK+TY, W-MO_SK+ST, and W-MO_SK+TY.

Table 2
Specification for Baseline Estimations

Abbreviation	Frequency Conversion	Sample Period	Fixed Effects		
			SKU	Store	Store Type
D-ROL_SK+ST	None	Two-week staggered rolling sample period	✓	✓	
D-ROL_SK+TY			✓		✓
W-MO_SK+ST	Weekly Mode	Whole sample period	✓	✓	
W-MO_SK+TY			✓		✓

The four specifications are a combination of strategies to deal with the two factors—the temporary sales effects and the retail service quality differences. The first part of the abbreviations, D-ROL or W-MO, corresponds to the strategy to deal with the effects of structural changes in temporary sales. D-ROL employs two-week staggered rolling estimations for daily frequency data with weekly time dummies. Rolling estimations are expected to absorb item-wide structural changes, including the frequency and price discount of temporary sales, by changes in constant terms over time. W-MO converts daily frequency data to weekly frequency data using the weekly mode of unit prices. The weekly mode conversion is a direct adjustment for temporary sales effects on unit prices, as discussed in the previous section.

The second part of the abbreviations, SK+ST or SK+TY, represents the fixed effects for a cross-sectional direction to control differences in the product quality and the retail service quality across the outlet channels. SK+ST employs two fixed effects for SKU code and store identity (ID) separately, thereby controlling the retail service quality differences at an individual store level. SK+TY employs SKU code and store type dummies for GMS, SM, M-SM, CVS, HC/DS,

[11] Shiratsuka (1999) pointed out that the CPI failed to reflect consumers' shift from retail shops and department stores to discount stores in the mid-1990s due to the limited coverage of discount stores in price surveys. Such structural changes in retail markets are called "price busting." It was induced by the massive expansion of large-scale shopping sites, such as general merchandise stores and large discount stores. After the sharp appreciation of the Japanese yen after the Plaza accord in 1985, it is often pointed out that the inefficiency of retail markets in Japan is the cause of price level differences between Japan and other advanced economies.

Drug/L, and Liquor/DS separately, thus, controlling the retail service quality differences at the outlet channel level. Comparing the estimation results for the two types of fixed effects enables us to confirm whether differences in retail service quality matter at a store or an outlet channel level. In addition, the second type of fixed effects also has the advantage that the estimates of fixed effects provide convenience premiums or discounts across outlet channels.

4.2 Estimation procedure

We construct the high-frequency quality-adjusted price indices in two steps. The first step is to estimate the extended TPD model for all 151 items, thereby computing elementary price indices for all items. The second step aggregates elementary indices to an overall index using sales amounts as weights.

As the first step for computing high-frequency quality-adjusted price indices, we estimate the extended TPD model for all 151 items based on the four baseline specifications, as shown in Table 2. The specifications for the daily frequency estimation of D-ROL_SK+ST and D-ROL_SK+TY are given by Eqs. (7) and (8), respectively.

$$lnUP_{i,k,t} = \lambda_w D\_WK^w_{i,k,t} + \gamma_i + \delta_k + \epsilon_{i,k,t}, \tag{7}$$

$$lnUP_{i,k,t} = \lambda_w D\_WK^w_{i,k,t} + \gamma_i + \mu_j + \epsilon_{i,k,t}, \tag{8}$$

$$D\_WK^w_{i,k,t} = \begin{cases} 1 & t \text{ belongs to the week } w \\ 0 & \text{otherwise} \end{cases},$$

where $UP_{i,k,t}$ and $D\_WK^w_{i,k,t}$ represent unit price and week dummy of second-week observations in the rolling subsample period for SKU code i at store ID k and time t, respectively. γ_i, δ_k, and μ_j are fixed effects for a product with SKU code i, a store with store ID k, and store type j, respectively. The fixed effect for product γ_i controls all the product characteristics, including the unit size, tracing a shadow price for the quality of a product i. The fixed effects for store δ_k and store type μ_j control all the quality differences related to retail services at store and store type levels, respectively. Here store ID k is included in store type j, i.e., GMS, SM, M-SM, CVS, HC/DS, Drug/L, and Liquor/DS. The estimated coefficient for week dummies (λ_w) corresponds to the quality-adjusted price index using week one as the base period.

The specifications for the weekly frequency estimation of W-MO_SK+ST and W-MO_SK+TY are given by Eqs. (9) and (10) below:

$$lnUP^{MO}_{i,k,w} = \sum_{v=2}^{W} \lambda_v D\_WK^v_{i,k,w} + \gamma_i + \delta_k + \epsilon_{i,k,w}, \tag{9}$$

$$lnUP^{MO}_{i,k,w} = \sum_{w=2}^{W} \lambda_v D\_WK^v_{i,k,w} + \gamma_i + \mu_j + \epsilon_{i,k,w}, \tag{10}$$

$$D\_WK^{v}_{i,k,t} = \begin{cases} 1 & t \text{ belongs to the week } v \\ 0 & \text{otherwise} \end{cases},$$

where $UP^{MO}_{i,k,w}$ and $D\_WK^{v}_{i,k,t}$ indicate the weekly mode of unit price and week dummies for SKU code i, at store ID k, and week v, respectively. γ_i, δ_k, and μ_j indicate the fixed effects for a product with SKU code i, a store with store ID k, and store type j, which are the same as the daily frequency estimation specifications.

Table 3 summarizes the estimation results for W-MO_SK+ST and W-MO_SK+TY on the three selected items—sweetened buns with the largest number of SKUs, honey with the median number of SKUs, and skimmed milk with the smallest number of SKUs. We find two important points about the estimation results in the table. First, the adjusted R-squared for all the estimations is generally high, indicating that the two types of fixed effects for a cross-sectional direction—SKU and store as well as SKU and store type—effectively control the time-invariant factors for product characteristics and retail service quality. Second, the estimated quality-adjusted price indices differ from product to product, even with generally high precision in the estimation results. The price indices for honey reveal statistically significant increases from the first week of January 2020 during most of the sample period, whereas those for skimmed milk remain almost unchanged, as seen in the statistically insignificant estimates for most of the weeks, and those for sweetened buns exhibit statistically significant decreases during the second half of the sample period.

We repeat the full sample estimation of the extended TPD model of Eqs. (9) and (10) for all the 151 items for the weekly frequency estimations. For the daily frequency estimation, we first conduct the two-week staggered rolling regressions on the extended TPD model of Eqs. (7) and (8) 24 times over the sample period of 25 weeks and then repeat that process for all the 151 items. This is the first step of computing the elementary price indices of the items.

In the second step, we aggregate elementary price indices to the overall price index by using sales amounts as weights. In the daily frequency estimation for the two-week subsample periods, the estimated coefficients for week dummies correspond to quality-adjusted price changes from the previous week. The cumulative quality-adjusted price changes are computed by linking the estimated week-on-week price changes from the base week. In the weekly frequency estimation for the full sample period, the estimated coefficients for week dummies correspond to the cumulative quality-adjusted price changes from the base week. Thus, the high-frequency quality-adjusted price indices for the week w in the daily and weekly frequency estimations—$HFQAPI^{D-ROL}_{w}$ and $HFQAPI^{W-MO}_{w}$—are computed as follows:

$$HFQAPI^{D-ROL}_{w} = \sum_{v=2}^{w} \sum_{i=1}^{I} sw_i \lambda_{v,i}, \tag{11}$$

$$HFQAPI^{W-MO}_{w} = \sum_{i=1}^{I} sw_i \lambda_{w,i}, \tag{12}$$

Table 3
Representative Estimation Results for Weekly Frequency Estimations

	Sweetened Buns (Largest number of SKUs)				Honey (Median number of SKUs)				Skimmed Milk (Smallest number of SKUs)			
	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE
2wk_id	−0.003 ***	(0.001)	−0.003***	(0.001)	0.005*	(0.003)	0.004	(0.003)	0.000	(0.004)	−0.001	(0.006)
3wk_id	0.001	(0.001)	0.001	(0.001)	−0.010***	(0.003)	−0.010***	(0.003)	−0.009*	(0.005)	−0.009	(0.007)
4wk_id	0.002 *	(0.001)	0.002*	(0.001)	0.001	(0.003)	0.002	(0.003)	−0.003	(0.004)	−0.002	(0.006)
5wk_id	0.000	(0.001)	0.000	(0.001)	0.010***	(0.003)	0.010***	(0.003)	−0.003	(0.005)	−0.004	(0.007)
6wk_id	−0.001	(0.001)	−0.001	(0.001)	−0.013***	(0.003)	−0.013***	(0.003)	−0.002	(0.005)	−0.004	(0.007)
7wk_id	0.001	(0.001)	0.001	(0.001)	0.009***	(0.003)	0.010***	(0.003)	0.000	(0.004)	0.000	(0.006)
8wk_id	0.002 **	(0.001)	0.002*	(0.001)	0.007***	(0.003)	0.008***	(0.003)	0.001	(0.004)	0.001	(0.006)
9wk_id	−0.002 **	(0.001)	−0.003**	(0.001)	0.011***	(0.003)	0.011***	(0.003)	−0.001	(0.004)	−0.001	(0.007)
10wk_id	−0.001	(0.001)	−0.002	(0.001)	0.015***	(0.003)	0.015***	(0.003)	−0.006	(0.006)	−0.002	(0.008)
11wk_id	−0.003 **	(0.001)	−0.003***	(0.001)	0.001	(0.003)	0.001	(0.003)	−0.008	(0.006)	−0.007	(0.007)
12wk_id	0.005 ***	(0.001)	0.004***	(0.001)	0.014***	(0.003)	0.013***	(0.003)	−0.004	(0.005)	−0.002	(0.007)
13wk_id	−0.001	(0.001)	−0.001	(0.001)	0.013***	(0.003)	0.012***	(0.003)	−0.007	(0.005)	−0.004	(0.007)
14wk_id	0.001	(0.001)	0.000	(0.001)	0.018***	(0.003)	0.017***	(0.003)	−0.001	(0.004)	0.002	(0.007)
15wk_id	−0.001	(0.001)	−0.002	(0.001)	0.017***	(0.003)	0.016***	(0.003)	−0.003	(0.005)	0.002	(0.007)
16wk_id	−0.002	(0.001)	−0.002**	(0.001)	0.011***	(0.003)	0.012***	(0.003)	−0.002	(0.005)	0.000	(0.007)
17wk_id	−0.005 ***	(0.001)	−0.005***	(0.001)	0.016***	(0.003)	0.017***	(0.003)	0.000	(0.004)	0.003	(0.006)
18wk_id	−0.009 ***	(0.001)	−0.009***	(0.001)	0.018***	(0.002)	0.018***	(0.003)	0.001	(0.004)	0.006	(0.006)
19wk_id	−0.004 ***	(0.001)	−0.005***	(0.001)	0.016***	(0.003)	0.017***	(0.003)	0.001	(0.003)	0.007	(0.006)
20wk_id	−0.002	(0.001)	−0.002	(0.001)	0.014***	(0.003)	0.014***	(0.003)	0.000	(0.004)	0.003	(0.006)
21wk_id	−0.004 ***	(0.001)	−0.005***	(0.001)	0.015***	(0.003)	0.015***	(0.003)	−0.001	(0.004)	−0.003	(0.006)
22wk_id	−0.010 ***	(0.001)	−0.010***	(0.001)	0.011***	(0.003)	0.011***	(0.003)	−0.003	(0.004)	−0.007	(0.008)
23wk_id	−0.007 ***	(0.001)	−0.007***	(0.001)	0.012***	(0.003)	0.013***	(0.003)	−0.002	(0.004)	−0.002	(0.006)
24wk_id	−0.006 ***	(0.001)	−0.006***	(0.001)	0.014***	(0.003)	0.015***	(0.003)	−0.001	(0.004)	0.001	(0.006)
25wk_id	−0.006 ***	(0.001)	−0.006***	(0.001)	0.013***	(0.003)	0.013***	(0.003)	−0.002	(0.004)	−0.002	(0.006)
Constant	4.744 ***	(0.001)	4.744***	(0.001)	0.737***	(0.002)	0.737***	(0.002)	0.728***	(0.003)	0.727***	(0.004)
Observations	449,013		449,013		36,414		36,414		4,111		4,111	
Adj. R-squared	0.882		0.870		0.988		0.986		0.965		0.931	
Fixed effects	SKU & Store		SKU & Store Type		SKU & Store		SKU & Store Type		SKU & Store		SKU & Store Type	

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

where sw_i is the sales amounts weight for item i, and I is the number of SKUs of the item. We use the fixed weight in this paper because of the two reasons. First, using the fixed or time-varying weights in the daily frequency estimations produces similar results over time. Second, the weekly frequency estimations with full sample data do not fit well with the time-varying weights aggregation.

4.3 Baseline estimation results

Fig. 9 summarizes the estimated HFQAPIs based on the baseline estimation results for D-ROL_SK+ST, D-ROL_SK+TY, W-MO_SK+ST, and W-MO_SK+TY. The plotted figures are the cumulative price changes on a quality-adjusted basis, using the week starting on January 6, 2020, as the base period. The figure also includes the UPI for reference. The light blue shaded area in the figures indicates the period of the first declaration of a state of emergency in Tokyo.

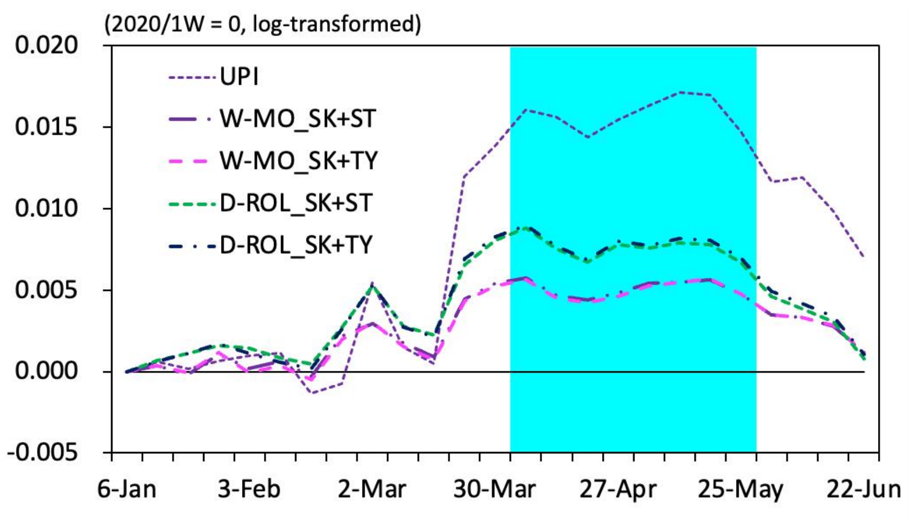


Fig. 9. Baseline Estimation Results
Notes: The light blue area indicates the period for the first declaration of a state of emergency in Tokyo.

The figure depicts three points regarding the inflation developments for food and beverage products during the first half of 2020. First, all the indicators, including UPI, reveal very similar trends until the end of February. Second, UPI and other quality-adjusted price indices reveal different trends since then because of the heightened concern over the wide and rapid spread of COVID-19 in Tokyo. Third, again, all the indicators reveal a similar downward trend from the end of the first declaration of a state of emergency.

The cumulative inflation from the beginning of January to the peak is estimated at 0.9% in the daily frequency estimations and 0.6% in the weekly frequency estimations. The estimates are slightly different but stay between UPI and zero inflation. In addition, the estimates for the daily and weekly specifications are not influenced by the two types of fixed effects in a cross-sectional direction, SKU and store as well as SKU and store type. This suggests that the retail service quality differences are well controlled at the store type level.

Fig. 10 plots the aggregated fixed effects of store types, δ, over all items using sales amount share as weights. The base store type is GMS, and the plotted figures indicate the convenience premiums or discounts of the store type against GMS. The estimates reveal very stable movements over time for all store types. This suggests that convenience premium or discount for store types remained unchanged under the first wave of the COVID-19 spread. CVS has the largest premium, reflecting its high retail service quality in terms of operating hours, location, and shopping time. SM and M-SM have smaller but positive premiums. HC/DS, Drug/L, and Liquor/DS give discounts, which is consistent with the sales strategy of lowering prices by reducing retail services. Discounts for Liquor/DS started shrinking slightly from March, reflecting the growing demand for home drinking under a voluntary lockdown.

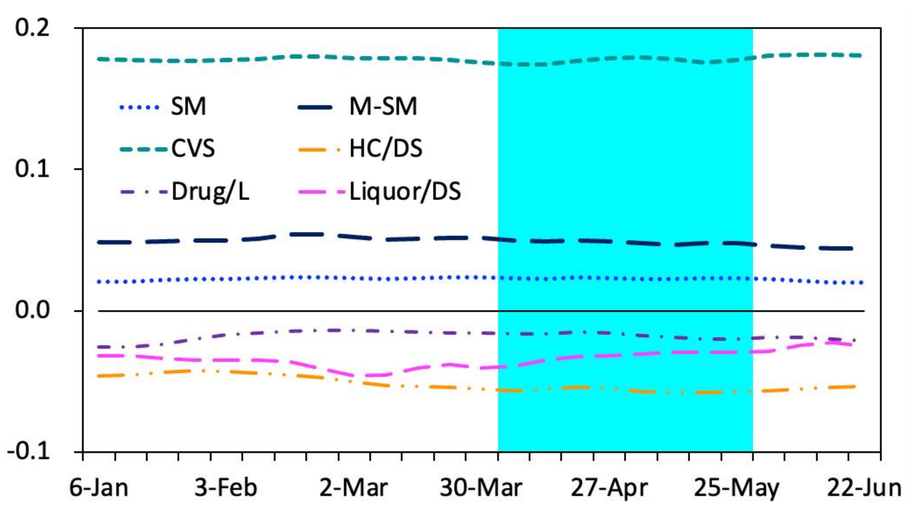


Fig. 10. Estimates for Store Type Fixed Effect
Notes: GMS is used as the base of the fixed effect on the store type, and the plotted figures indicate the convenience premiums or discounts of the store type against GMS. The light blue area indicates the period for the first declaration of a state of emergency in Tokyo.

The results of the baseline estimations seem consistent with Fig. 1, which indicates that the inflation of UPI (CPINow T-index) accelerated, whereas there was no acceleration in the mode UPI (CPINow T-mode-index). The T-index fully reflects the decrease in temporary sales as a price increase, whereas the T-mode-index does not account for changes in temporary sales. The quality-adjusted price indices, in terms of both product characteristics and retail service quality, are somewhere between the two CPINow indicators because households try to minimize the risk of COVID-19 infection by reducing travel distance to outlets and shopping time. In the next section, we examine the appropriateness of the estimated quality-adjusted price index with various robustness checks.

5 Robustness Checks

We now carry out the robustness checks from the following five perspectives: (1) alternative fixed effect for product-store combination, (2) whole sample estimation using daily frequency data, (3) alternative summary statistics for weekly conversion (mean and median), (4) inclusion of temporary sales dummy as an explanatory variable for the daily frequency estimation, and (5) extension of subsample period to three weeks in the staggered rolling regression. Table 4 summarizes the five robustness checks.

Table 4
Robustness Checks

<table>
<tr><th></th><th>Abbreviations</th><th>Details on Robustness Checks</th></tr>
<tr><td rowspan="2">(1)</td><td>D-ROL_SKxST</td><td rowspan="2">Alternative fixed effect for product-store combination (SKxST)</td></tr>
<tr><td>W-MO_SKxST</td></tr>
<tr><td rowspan="2">(2)</td><td>D-WHL_SK+ST</td><td rowspan="2">Whole sample estimation using daily frequency data (WHL)</td></tr>
<tr><td>D-WHL_SK+TY</td></tr>
<tr><td rowspan="4">(3)</td><td>W-MN_SK+ST</td><td rowspan="4">Alternative summary statistics for weekly conversion: mean (MN) and median (MD)</td></tr>
<tr><td>W-MN_SK+TY</td></tr>
<tr><td>W-MD_SK+ST</td></tr>
<tr><td>W-MD_SK+TY</td></tr>
<tr><td rowspan="3">(4)</td><td>D-ROL_SK+ST_TS</td><td rowspan="3">Inclusion of temporary sales dummy as an explanatory variable for daily frequency estimation (TS)</td></tr>
<tr><td>D-ROL_SK+TY_TS</td></tr>
<tr><td>D-ROL_SKxST_TS</td></tr>
<tr><td rowspan="2">(5)</td><td>D-ROL3_ST+TY_1</td><td rowspan="2">Extension of subsample period to three weeks in staggered rolling regression with different linking methods</td></tr>
<tr><td>D-ROL3_ST+TY_2</td></tr>
</table>

5.1 Robustness check 1: Fixed effect for the product–store combination

The first robustness check examines the impacts of alternative fixed effects for the product-store combination as the cross term of SKU code and store ID. In this case, the same products sold at different stores are considered different goods. The estimation specifications of D-ROL_SKxST and W-MO_SKxST are given by Eqs. (13) and (14) for the daily and weekly frequency estimations, respectively:

$$lnUP_{i,k,t} = \lambda_w D\_WK^{w}_{i,k,t} + \eta_{i,k} + \epsilon_{i,k,t}, \tag{13}$$

$$lnUP^{MO}_{i,k,w} = \sum_{v=2}^{W} \lambda_v D\_WK^{v}_{i,k,w} + \eta_{i,k} + \epsilon_{i,k,w}, \tag{14}$$

where $\eta_{i,k}$ is a fixed effect for the product–store combination for SKU code i and store ID k.

Fig. 11 plots D-ROL_SKxST and W-MO_SKxST along with the baseline specifications of D-ROL_SK+ST and W-MO_SK+ST as well as the UPI. Conducting the estimations with the cross term of SKU and store as a fixed effect for a cross-sectional direction produces lower estimates for the quality-adjusted price indices in both the daily and weekly frequency estimations. The downward revision is larger in the daily frequency estimation. This estimation results suggest the possibility of the overadjustments of SKU-store specific price increases when using the cross term of SKU and store as a fixed effect. This tendency is further amplified by the two-week staggered rolling estimations for daily frequency data through the time-varying constant terms.

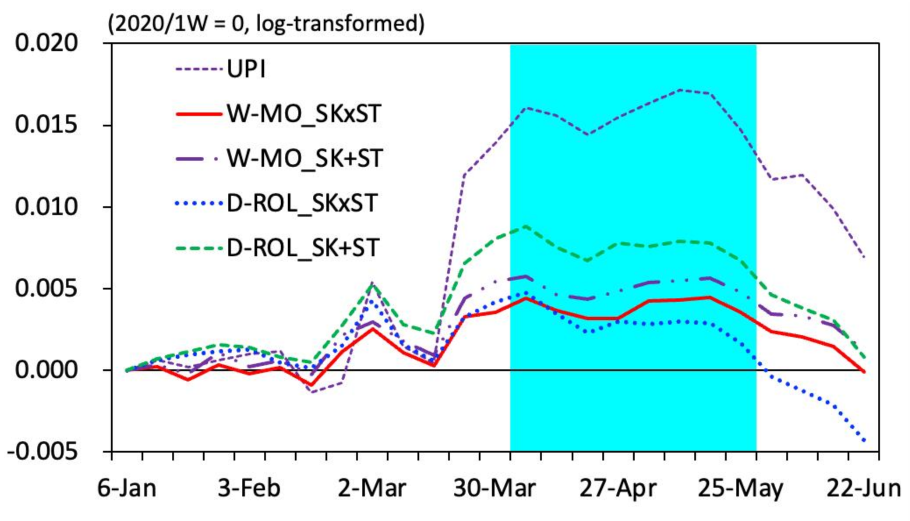


Fig. 11. Robustness Check 1: Fixed Effect for the Product-Store Combination
Notes: The light blue area indicates the period of the first declaration of a state of emergency in Tokyo.

5.2 Robustness check 2: Daily estimation for the full sample period

The second robustness check focuses on the effects of structural changes in temporary sales in the daily frequency estimations by comparing the results of the daily estimation for the full sample period data with those of the baseline two-week staggered rolling estimations. The staggered rolling estimations are expected to absorb item-wide structural changes, including the frequency and price discount of temporary sales, by changes in constant terms over time. The estimation specifications of D-WHL_SK+ST and D-WHL_SK+ST are given by Eqs. (15) and (16) for the two types of fixed effects, respectively:

$$lnUP_{i,k,t} = \sum_{w=2}^{W} \lambda_w D\_WK^{w}_{i,k,t} + \gamma_i + \delta_k + \epsilon_{i,k,t}, \tag{15}$$

$$lnUP_{i,k,t} = \sum_{w=2}^{W} \lambda_w D\_WK^{w}_{i,k,t} + \gamma_i + \mu_j + \epsilon_{i,k,t}, \tag{16}$$

Fig. 12 depicts the estimation results for D-WHL_SK+ST and D-WHL_SK+TY, along with the baseline estimation results for D-ROL_SK+ST and D-ROL_SK+TY as well as the UPI. The cumulative inflation for D-WHL_SK+ST and D-WHL_SK+TY are both estimated at 1.2%, which is higher than the baseline estimation results. However, the estimates are not influenced by the two types of fixed effects in a cross-sectional direction for SKU and store as well as SKU and store type.

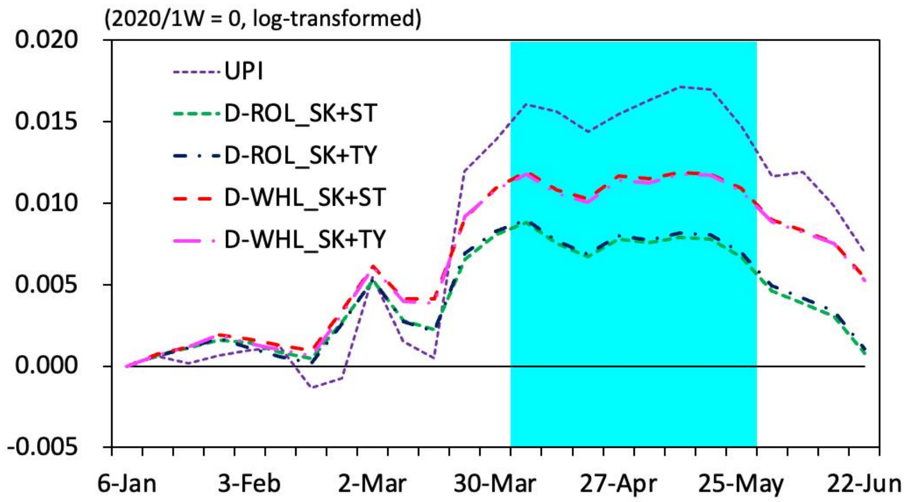


Fig. 12. Robustness Check 2: Daily Estimation for the Full Sample Period
Notes: The light blue area indicates the period of the first declaration of a state of emergency in Tokyo.

The differences between D-ROL and D-WHL are estimated at 0.3%. The differences come from the time-varying constant terms in the two-week staggered rolling estimations. The time-varying constant terms are deemed effective in absorbing item-wide structural changes, including the frequency and price discount of temporary sales. Thus, we regard the differences as the effects of reduced frequency and price reduction of temporary sales in the daily frequency estimations. In addition, the differences between UPI and D-WHL of around 0.5% are regarded as outlet substitution effects stemming from the households' behavior to purchase from a store nearby at slightly higher prices because of the risk of COVID-19 infection.

5.3 Robustness check 3: Weekly data conversion measures

The third robustness check focuses on the effects of structural changes in temporary sales in the weekly frequency estimations by comparing the estimation results using two alternative summary statistics for data conversion on a weekly basis—mean (MN) and median (MD). The estimation specifications of W-MN_SKxST, W-MN_SKxST, W-MN_SKxST, and W-MN_SKxST are given by Eqs. (17), (18), (19), and (20) for the two types of fixed effects, respectively:

$$lnUP^{MN}_{i,k,w} = \sum_{v=2}^{W} \lambda_v D\_WK^{v}_{i,k,w} + \gamma_i + \delta_k + \epsilon_{i,k,w}, \tag{17}$$

$$lnUP^{MN}_{i,k,w} = \sum_{w=2}^{W} \lambda_v D\_WK^{v}_{i,k,w} + \gamma_i + \mu_j + \epsilon_{i,k,w}, \tag{18}$$

$$lnUP^{MD}_{i,k,w} = \sum_{v=2}^{W} \lambda_v D\_WK^{v}_{i,k,w} + \gamma_i + \delta_k + \epsilon_{i,k,w}, \tag{19}$$

$$lnUP^{MD}_{i,k,w} = \sum_{w=2}^{W} \lambda_v D\_WK^{v}_{i,k,w} + \gamma_i + \mu_j + \epsilon_{i,k,w}. \tag{20}$$

The estimation results in Fig. 13 are ordered from highest to lowest as follows: W-MN, W-MD, and W-MO, with almost identical estimates for the two types of fixed effects in a cross-sectional direction. The results suggest that W-MN is mostly influenced by declines in the frequency and size of price reductions of temporary sales. This is consistent with the observation in Fig. 7 that compares the inflation trend with the three statistical measures for the central tendency of unit price distributions, revealing that the mean is the highest, followed by the mean and median. Thus, it is likely that W-MN and W-MD make overadjustments of the temporary sales effects.

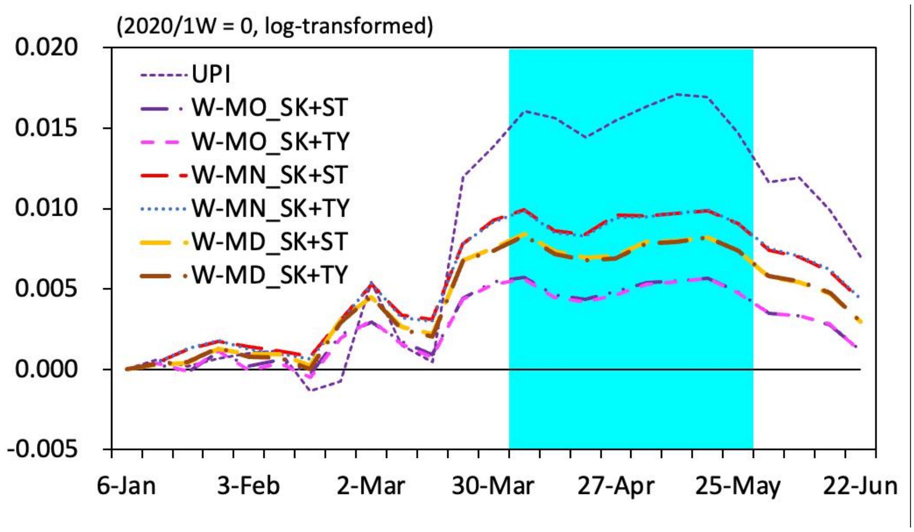


Fig. 13. Robustness Check 3: Weekly Data Conversion Measures
Notes: The light blue area indicates the period of the first declaration of a state of emergency in Tokyo.

As in the previous robustness check for the daily full sample estimations, the differences between W-MN and W-MO (0.4%) and those between UPI and W-MN (0.7%) are regarded as the effects of temporary sales and outlet substitution, respectively. The differences between W-MN and W-MO solely come from the summary statistics measures in weekly frequency conversion, that is, the mean or mode of unit prices in a week, which corresponds to the effects of temporary sales in terms of reduced frequency and price reduction.

5.4 Robustness check 4: Estimation with the temporary sales dummy

The fourth robustness check introduces the temporary sales dummy as an additional explanatory variable in the two-week staggered rolling estimations with daily data. This robustness check examines the existence of additional temporary sales effects, which are not detected in the baseline estimation specifications for the daily frequency estimations. As explained in Section 3, temporary sales are defined as a day when the observed price for a product at a store is at least two yen lower than the mode price for the past seven days, including the current day, following Abe and Tonogi (2010).

The estimation specifications of D-ROL_SK+ST_TS, D-ROL_SK+TY_TS, and D-ROL_SKxST_TS are given by Eqs. (21), (22), and (23).

$$lnUP_{i,k,t} = \lambda_w D\_WK^w_{i,k,t} + \kappa D\_TS_{i,k,t} + \gamma_i + \delta_k + \epsilon_{i,k,t}, \tag{21}$$

$$lnUP_{i,k,t} = \lambda_w D\_WK^w_{i,k,t} + \kappa D\_TS_{i,k,t} + \gamma_i + \mu_j + \epsilon_{i,k,t}, \tag{22}$$

$$lnUP_{i,k,t} = \lambda_w D\_WK^w_{i,k,t} + \kappa D\_TS_{i,k,t} + \eta_{i,k} + \epsilon_{i,k,t}, \tag{23}$$

$$D\_TS_{i,k,t} = \begin{cases} 1 & \text{temporary sales for SKU code } i\text{, store id } k\text{, at time } t \\ 0 & \text{otherwise} \end{cases}$$

where $D\_TS_{i,k,t}$ indicate the temporary sales dummy for SKU code i at store ID k and time t. κ is the estimated coefficient for the temporary sales dummy.

Fig. 14 depicts the estimation results. Surprisingly, all the three specifications with the temporary sales dummy—D-ROL_SK+ST_TS, D-ROL_SK+TY_TS, and D-ROL_SKxST_TS—exhibit almost identical fluctuations over time, suggesting that the inclusion of the temporary sales dummy eliminates the temporary price fluctuation at the individual price observation level. However, the result is consistent with the no acceleration in the year-on-year changes in the CPINow T-mode-index, as depicted in Fig. 1. The above observations suggest that introducing the temporary sales dummy as an additional explanatory variable leads to overadjustments of the temporary sales effects by eliminating all the temporary price reductions.

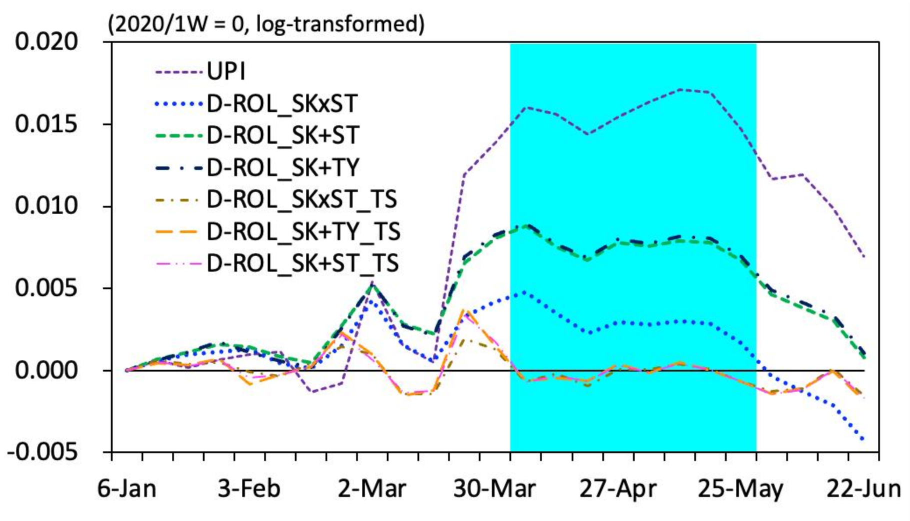


Fig. 14. Robustness Check 4: Estimation with Temporary Sales Dummies
Notes: The light blue area indicates the period of the first declaration of a state of emergency in Tokyo.

5.5 Robustness check 5: three-week staggered rolling regression

We perform the final robustness check by extending the subsample periods for the staggered rolling regression from two to three weeks, as given by Eq. (24).

$$lnUP_{i,k,t} = \lambda_w D\_WK^{w}_{i,k,t} + \lambda_{w+1} D\_WK^{w+1}_{i,k,t} + \gamma_i + \mu_j + \epsilon_{i,k,t}. \quad (24)$$

We construct two types of quality-adjusted price indices based on the estimation results using the estimated coefficients for the second- and third-week dummies, denoted as D-ROL3+SK+TY_1 and D-ROL3+SK+TY_2, respectively. The first linking method, D-ROL3+SK+TY_1, links the most recent period movement to the previous estimates. The second linking method, D-ROL3+SK+TY_2, links the whole period movement to the previous estimates.

As depicted in Fig. 15, D-ROL3+SK+TY_1 and D-ROL3+SK+TY_2 move almost identically and very close to D-ROL_SK+TY. This confirms that extending the subsample period to three weeks does not significantly influence the estimation results.

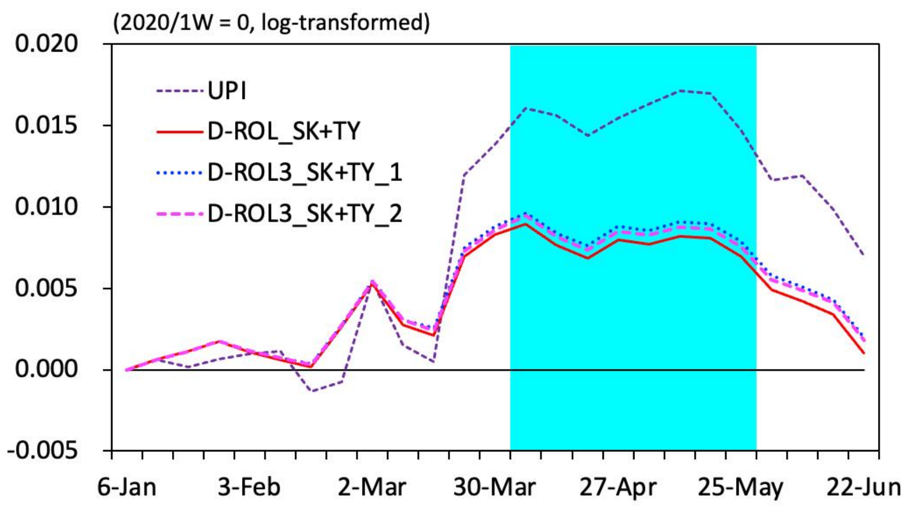


Fig. 15. Robustness Check 5: Three-Week Staggered Rolling Estimations
Notes: The light blue area indicates the period for the first declaration of the state of emergency in Tokyo.

6 Inflation Trend Under the First Wave of the COVID-19 Spread

Table 5 decomposes changes in UPI into three components—temporary sales effects, outlet substitution effects, and changes in quality-adjusted price indices. In the daily frequency estimation, the increase in UPI of 1.7% is decomposed into the temporary sales effects of 0.3%, the outlet substitution effects of 0.5%, and the increase in the quality-adjusted price index of 0.9%, whereas in the weekly frequency estimation, it is 0.4%, 0.7%, and 0.6%, respectively. The observed increases in UPI overestimate the quality-adjusted price increases caused by the inappropriate adjustments of the outlet substitution effects and the temporary sales effects.

The cumulative increases in quality-adjusted price indices reached their peaks at 0.9% in the daily frequency estimation and 0.6% in the weekly frequency estimation in early April, remaining around those levels until late May and then started declining. Due to the possibility of different assessments of the temporary sales effects from person to person, inflation is 1.2% in the daily frequency estimation and 1.1% in the weekly frequency estimation by counting the

Table 5
Decomposition of Unit Price Changes

	Daily Estimation	Weekly Estimation
Unit Price Index	1.7% (= UPI)	
Outlet Substitution Effects	0.5% (= UPI − D-WHL)	0.7% (= UPI − W-MN)
Temporary Sales Effects	0.3% (= D-WHL − D-ROL)	0.4% (= W-MN − W-MO)
Quality-adjusted Price Index	0.9% (= D-ROL)	0.6% (= W-MO)

temporary sales effects as inflation.

Our answer to the question of this paper, "Was inflation observed under the first wave of the COVID-19 spread?" is yes, but it was fairly mild and temporary.

7 Conclusions

We examined whether inflation was observed under the first wave of the COVID-19 spread in Japan. To address this question, we constructed high-frequency quality-adjusted price indices by employing daily scanner data of retail stores in Tokyo. We made explicit adjustments for not only the product characteristics but also the structural changes in temporary sales and the retail service quality of outlet channels. We assumed that adjustments for the effects of temporary sales and retail service quality are particularly important in examining the retail price dynamics under the COVID-19 pandemic as the voluntary lockdown constrained household purchasing behavior.

We concluded that mild and temporary inflation of slightly less than 1% (0.6%-0.9%) was observed during the first wave of the COVID-19 spread based on the overall assessment of the baseline estimations and the five robustness checks. We also decomposed the differences between the increases in the unit prices and those in the quality-adjusted price indices into outlet substitution effects and temporary sales effects. In the daily frequency estimations, the increase in UPI of 1.7% was decomposed into the temporary sales effects of 0.3%, the outlet substitution effects of 0.5%, and the increases in the quality-adjusted price indices of 0.9%. In the weekly frequency estimations, it was 0.4%, 0.7%, and 0.6%, respectively.

Our empirical results revealed that the construction of quality-adjusted price indices becomes very difficult when faced with large-scale structural changes in the retail markets, including the frequency of temporary sales and the size of price reductions. As discussed in Section 1, the widening deviations between the two types of daily price indices—the T-index and the T-mode-index—indicate the possibility of structural changes in the retail markets. Fig. 16 plots the long time series of the deviations from 1990. As depicted in the figure, the period of our focus, January to June 2020, experienced the largest positive deviations between the two indices, suggesting that structural changes in retail markets were unprecedentedly large, although for a short period.[12]

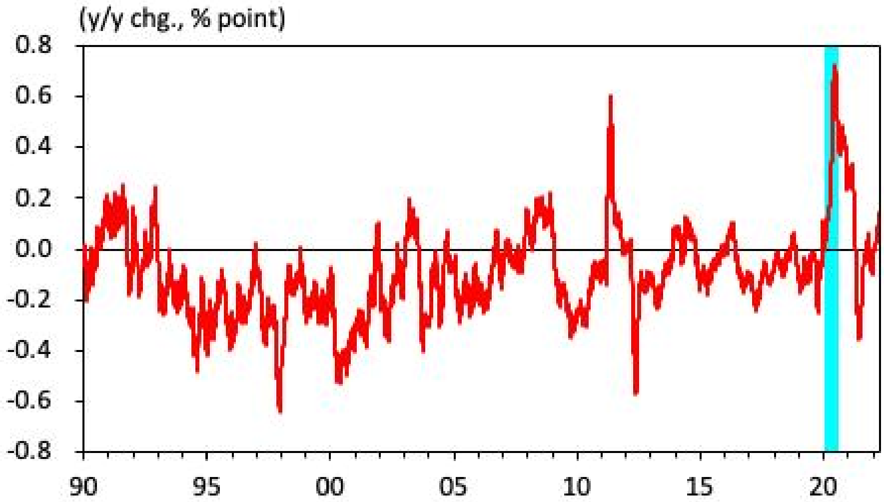


Fig. 16. Spreads between T-index and T-mode-index
Notes: The figure plots the divergence between year-on-year changes in T-index and T-mode-index. The light blue area indicates the first half of 2020, that is, the period under primary analysis in this paper.
Source: Nowcast, *CPINow*.

However, it should be noted that the mid-1990s to the early 2000s also experienced significant and continued deviations between the two indices in a negative direction. This period corresponds to observing the "price busting" phenomenon when large-scale structural changes in the retail markets, symbolized by the expansion of large-scale retail stores, were advanced. As Shiratsuka (1999) has pointed out, the outlet substitution bias severely affected the Japanese CPI during that period.

We stress that the CPI needs to consider how to deal with possible large-scale structural changes in the retail markets as it is hard to anticipate such changes in advance. It is also important to respond flexibly to such structural changes in the retail markets once it occurs. However, the current price survey method based on a "one-specification-for-one-item" policy is unable to deal with that issue, as the method specifies a few most popular specifications for each item and continuously surveys their prices at specific outlets.[13] The most promising alternative price information is the use of scanner data, as considered in this paper. Although scanner data cover about 20% of the CPI basket, these products represent a significant portion of the Retail Price Survey, which is the primary data source for the CPI. It is a future challenge to explore an estimation framework to produce stable results while flexibly incorporating the effects of structural changes with longer-term scanner data.

[12] Another major divergence was observed immediately after the Great East Japan Earthquake in March 2011.

[13] Shiratsuka (2021) examined the lower-level substitution bias using the micro-data for the Retail Price Survey, which is the primary source data for the Japanese CPI. He revealed that the lower-level substitution bias stemming from the elementary aggregation formula is very limited but volatile in both positive and negative directions.

References

Abe, N., Inoue, T., Sato, H. 2020. Price Index Numbers Under Large-Scale Demand Shocks: The Japanese Experience of the COVID-19 Pandemic. RCESR Discussion Paper Series DP20-2, Research Center for Economic and Social Risks, Institute of Economic Research, Hitotsubashi University.

Abe, N., Tonogi, A. 2010. Micro and macro price dynamics in daily data. Journal of Monetary Economics. 57 (6), 716–728.

Cavallo, A. 2020. Inflation with covid consumption baskets. Working Paper 27352, National Bureau of Economic Research.

Diewert, W. E., Fox, K. J. 2020. Measuring real consumption and cpi bias under lockdown conditions. Working Paper 27144, National Bureau of Economic Research.

Eichenbaum, M., Jaimovich, N., Rebelo, S. 2011. Reference prices, costs, and nominal rigidities. American Economic Review. 101 (1), 234–62.

Fujii, D., Nakata, T. 2021. Covid-19 and output in japan. The Japanese Economic Review. 72 (4), 609–650.

de Haan, J. 2015. A framework for large scale use of scanner data in the dutch cpi. paper presented at the fourteenth ottawa group meeting, tokyo, japan.

de Haan, J., Hendriks, R. 2013. Price measurement using scanner data: Time-product dummy versus time dummy hedonic indexes. paper presented at the 2013 economic measurement group, sydney, australia.

de Haan, J., Hendriks, R., Scholz, M. 2021. Price measurement using scanner data: Time-product dummy versus time dummy hedonic indexes. Review of Income and Wealth. 67 (2), 394–417.

de Haan, J., Krsinich, F. 2014. Scanner data and the treatment of quality change in nonrevisable price indexes. Journal of Business and Economic Statistics. 32 (3), 341–358.

Hosono, K. 2021. Epidemic and economic consequences of voluntary and request-based lockdowns in japan. Journal of the Japanese and International Economies. 61, 101147.

Imai, S., Watanabe, T. 2014. Product Downsizing and Hidden Price Increases: Evidence from Japan's Deflationary Period. Asian Economic Policy Review. 9 (1), 69–89.

Kehoe, P., Midrigan, V. 2015. Prices are sticky after all. Journal of Monetary Economics. 75, 35–53.

Konishi, Y., Saito, T., Ishikawa, T., Kanai, H., Igei, N. 2021. How Did Japan Cope with COVID-19? Big Data and Purchasing Behavior. Asian Economic Papers. 20 (1), 146–167.

Krsinich, F. 2016. The fews index: Fixed effects with a window splice. Journal of Official Statistics. 32 (2), 375–404.

Nakamura, E., Steinsson, J. 2008. Five Facts about Prices: A Reevaluation of Menu Cost Models*. The Quarterly Journal of Economics. 123 (4), 1415–1464.

Shibamoto, M., Hayaki, S., Ogisu, Y. 2022. Covid-19 infection spread and human mobility. Journal of the Japanese and International Economies. 64, 101195.

Shiratsuka, S. 1995. Automobile Prices and Quality Changes : A Hedonic Price Analysis of Japanese Automobile Market. Monetary and Economic Studies. 13 (2), 1–44.

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

Shiratsuka, S. 2021. Lower-level substitution bias in the japanese consumer price index: Evidence from government micro data. Discussion Paper Series A No.722, Hitotsubashi University, Institute for Economic Research.

Sudo, N., Ueda, K., Watanabe, K., Watanabe, T. 2018. Working less and bargain hunting more: Macroimplications of sales during japan's lost decades. Journal of Money, Credit and Banking. 50 (2-3), 449–478.

Ueda, K., Watanabe, K., Watanabe, T. 2019. Product turnover and the cost-of-living index: Quality versus fashion effects. American Economic Journal: Macroeconomics. 11 (2), 310–47.

Watanabe, T., Omori, Y. 2020. How Much Did People Refrain from Service Consumption due to the Outbreak of COVID-19?. CARF F-Series CARF-F-477, Center for Advanced Research in Finance, Faculty of Economics, The University of Tokyo.

Watanabe, T., Yabu, T. 2021. Japan's voluntary lockdown: further evidence based on age-specific mobile location data.. Japanese economic review (Oxford, England). 72, 333–370.