TCER Working Paper Series

Scanner Data and the Construction of Inter-Regional Price Indexes

Chihiro Shimizu
Erwin Diewert
Naohito Abe
Akiyuki Tonogi

February 2025

Working Paper E-211
https://www.tcer.or.jp/wp/pdf/e211.pdf

TOKYO CENTER FOR ECONOMIC RESEARCH
1-7-10-703 Iidabashi, Chiyoda-ku, Tokyo 102-0072, Japan

Abstract

The paper uses monthly scanner data on purchases of rice in six Japanese Prefectures over the 24 months in the years 2021 and 2022 in order to calculate alternative price indexes that are free from chain drift. The paper also attempts to measure the welfare effects of differing product availability across the six prefectures. In order to eliminate the chain drift problem, the following multilateral indexes were computed: GEKS, Geary-Khamis and Weighted Time Product Dummy Hedonic price indexes. Chain drift can also be eliminated by estimating purchaser preferences using consumer demand theory. Thus the paper uses the Japanese rice data to estimate linear preferences, CES preferences and Konüs Byushgens Fisher preferences (with a rank one substitution matrix). Feenstra (1994) worked out a method for measuring the gains (or losses) of utility from new and disappearing products and his method is adapted to measuring the welfare effects of differing degrees of product availability across the Prefectures.

Chihiro Shimizu
TCER
and
Hitotsubashi University
School of Social Data Science
2-1 Naka, Kunitachi, Tokyo, 186-8601
c.shimizu@r.hit-u.ac.jp

Erwin Diewert
University of British Columbia
Vancouver School of Economics
6000 Iona Drive Vancouver, BC Canada V6T 1L4
erwin.diewert@ubc.ca

Naohito Abe
TCER
and
Hitotsubashi University
Institute of Economic Research
2-1 Naka, Kunitachi, Tokyo, 186-8601
nabe@ier.hit-u.ac.jp

Akiyuki Tonogi
Toyo University
Department of Economics
5-28-20 Hakusan, Bunkyo-ku, Tokyo, 112-8606
tonogi@toyo.jp

Scanner Data and the Construction of Inter-Regional Price Indexes[*]

W. Erwin Diewert[†], Naohito Abe[‡], Akiyuki Tonogi[§] and Chihiro Shimizu.[¶]

February 22, 2025

Abstract

The paper uses monthly scanner data on purchases of rice in six Japanese Prefectures over the 24 months in the years 2021 and 2022 in order to calculate alternative price indexes that are free from chain drift. The paper also attempts to measure the welfare effects of differing product availability across the six prefectures. In order to eliminate the chain drift problem, the following multilateral indexes were computed: GEKS, Geary-Khamis and Weighted Time Product Dummy Hedonic price indexes. Chain drift can also be eliminated by estimating purchaser preferences using consumer demand theory. Thus the paper uses the Japanese rice data to estimate linear preferences, CES preferences and Konüs Byushgens Fisher preferences (with a rank one substitution matrix). Feenstra (1994)[35] worked out a method for measuring the gains (or losses) of utility from new and disappearing products and his method is adapted to measuring the welfare effects of differing degrees of product availability across the Prefectures.

Keywords

Consumer demand theory, scanner data, chain drift, multilateral indexes, CES and KBF utility functions, hedonic regressions, inter-regional price and quantity indexes.

JEL Classification Numbers

C33, C43, C81, D12

* The authors thank Robert Allen, Jan de Haan, Robert Feenstra, Hideyuki Mizobuchi, Alicia Rambaldi and Prasada Rao for helpful comments on earlier versions of the paper. The authors gratefully acknowledge financial support from the JSPS KAKEN Grant Grant 23K25496 & 24H00012.

† University of British Columbia and University of New South Wales: Email: erwin.diewert@ubc.ca

‡ Hitotsubashi University; Email: nabe@ier.hit-u.ac.jp

§ Toyo University; Email: tonogi@toyo.jp

¶ Hitotsubashi University; Email: c.shimizu@r.hit-u.ac.jp

1 Introduction

The paper addresses two main questions:

- How can we construct interregional price (and quantity) indexes for a country at the first stage of aggregation that are transitive over time and space?
- How can we measure the effects on price levels and welfare of smaller choice sets for regions that have a limited availability of products?

We will use household price and quantity data on purchases of the 80 top selling rice products over 24 months for six Prefectures in Japan in order to provide possible answers to the above questions.*1 Many alternative multilateral indexes will be constructed and compared using this data set.

The transitivity problem can be explained as follows. A bilateral index number formula that provides an estimate of the price level in one region or period to the price level in another region or period is basically a weighted average of ratios of product prices where the price of each product n in one region-month is compared to the same product n price in another region-month. The weights for the product price ratios are typically an average of the monthly expenditure shares on the products in the two region-months under consideration. Suppose we want to compare the prices in period 3 with the same prices in period 1 for the same region. Then we could construct a *fixed base bilateral index number* that directly compared the period 3 prices to the period 1 prices. Call this index $P(3/1)$. Alternatively, we could do a series of comparisons, comparing the prices of period 2 to period 1, obtaining the index $P(2/1)$, and then comparing the prices of period 3 to the corresponding period 2 prices, obtaining the index $P(3/2)$. The *chained index* between periods 3 and 1 is the product of the two chain link indexes, $P(2/1)$ times $P(3/2)$. We would like $P(3/1)$ to equal $P(2/1) \times P(3/2)$ but this *path independence or transitivity property* frequently fails. When this property fails, we say that we have a *chain drift problem.**2 Szulc (1983)[65] (1987)[66] demonstrated how big the chain drift problem could be using chained Laspeyres indexes but at that time, it was thought that chained *superlative indexes**3 would exhibit minimal chain drift.*4 However, de Haan (2008)[12] using Dutch data on *weekly* sales of detergents showed that chained Fisher indexes declined to about 10% of the initial index level after 150 weeks of data.*5 For more recent

*1 Our paper is similar to the recent paper by Fox, Levell and O'Connell (2024)[37] who compared several multilateral methods using scanner data. It is also related to the papers by Diewert and Fox (2021)[30] and Melser and Webster (2021)[57] who used simulated data that was exact for CES preferences and made comparisons between many alternative multilateral indexes.

*2 Fisher (1922; 293)[36] realized that many chained indexes in use at the time could be subject to chain drift but for his empirical data, there was no clear evidence of chain drift for his Fisher ideal index number formula. However, Persons (1921; 110)[58] came up with an empirical example where the Fisher index exhibited substantial downward chain drift. Frisch (1936; 9)[38] seems to have been the first to use the term "chain drift". Both Frisch (1936; 8-9)[38] and Persons (1928; 100-105)[59] discussed and analyzed the chain drift problem.

*3 See Diewert (1976)[18] for his definition of a superlative index. Basically, a superlative bilateral index number formula gives the "right" answer if consumers are maximizing an associated utility function that that can approximate an arbitrary linearly homogeneous utility function to the second order around any given point. Superlative index number formulae (like the Fisher (1922)[36] ideal index) are thought to deal with substitution effects in a satisfactory manner. The Laspeyres, Paasche and Fisher and other bilateral index number formulae will be defined below.

*4 See Hill (1988; 136-137)[50].

*5 The problem is caused by huge fluctuations in the volume sold when products are sold at a discount; see Diewert (2023)[25] on this point. For more information on the Dutch experiments with scanner data, see de Haan and van der Grient (2011)[14].

examples of massive downward chain drift of *monthly* chained superlative Törnqvist indexes using scanner data, see Fox, Levell and O'Connell (2024)[37].

The chain drift problem was not a problem before 2008, because before scanner data became available to National Statistical Offices, consumer (and producer) price indexes were produced in a very different way. At the first stage of aggregation, a sample of prices in a particular product category was collected in each month and these prices were compared to the same prices in the base month and either the arithmetic or geometric average of these product prices was taken as an estimate of the average price level of the current month to the price level of the product category in the base month. A weighted average of these product category price level ratios was taken to approximate the national price level for the current month relative to the base month. Annual weights for the aggregate product categories for some base year were used at the final stages of aggregation to weight the category price indexes which did not use weights. The annual weights came from periodic household expenditure surveys. Thus the tremendous fluctuations in product quantities that are frequently observed at the first stage of aggregation played no role in this historic way of constructing consumer and producer price indexes and thus there was no chain drift problem in the periods prior to 2008.

The paper by de Haan (2008)[12] led researchers to look for solutions to the chain drift problem. Thus Ivancic, Diewert and Fox (2009)[51] (2011)[52] suggested using multilateral index number theory on a rolling window of observations to mitigate the chain drift problem.*6 This strategy was eventually implemented by the Australian Bureau of Statistics (2016)[2].

A problem with using rolling window multilateral methods is that as an extra period of data becomes available, the indexes have to be recomputed and linked to the previous indexes. But how exactly are the results from the new window to be linked to the previous index values?*7 In this paper, we will not address this issue; it is a subject for ongoing research.*8

Section 2 of the paper provides some information on the data used in this study as well as introducing some notation that will be used throughout the paper. The remaining sections of the paper describe the various multilateral indexes that are used in this study. Our goal is to calculate various multilateral indexes using our Japanese panel data on sales of rice products for six Prefectures and the 24 months in the years 2021-2022. We will use multilateral indexes which are transitive, invariant to changes in the units of measurement and satisfy a strong identity test for quantities.

Section 3 describes our first multilateral method, the GEKS indexes. These indexes are based on matched product bilateral Fisher indexes so in this section, we also calculate fixed base Fisher indexes that take the first month in 2021 for Tokyo (the biggest Prefecture) as the base

*6 Balk (1980)[3] (1981)[4] suggested the use of multilateral indexes in the seasonal product context. More recently, Hill (2001)[48] (2004)[49] used multilateral indexes in the time series context. The origins of multilateral index number theory date back to Gini (1924)[42] (1931)[43]. For more recent reviews of multilateral methods, see Balk (1996)[5] (2008)[7] and Diewert (1999)[19] (2023)[25].

*7 Ivancic, Diewert and Fox (2011)[52] suggested that the movement of the rolling window indexes for the last two periods in the new window be linked to the last index value generated by the previous window. However Krsinich (2016)[55] suggested that the movement of the indexes generated by the new window be linked to the previous window index value for the second period in the previous window. Krsinich called this a *window splice* as opposed to the IDF *movement splice*. De Haan (2015; 27)[13] suggested that perhaps the linking period should be in the middle of the old window which the Australian Bureau of Statistics (2016; 12)[2] termed a *half splice*. Diewert and Fox (2021)[30] suggested that the *average* of all links for the last period in the new window to the observations in the old window could be used as the linking factor. Finally, Diewert and Shimizu (2024)[31] suggested an *ever expanding window* approach.

*8 For systematic treatments of the issues surrounding the implementation of a rolling window multilateral method and further discussion of extension methods, see the Australian Bureau of Statistics (2016)[2], Chessa (2016)[9] (2021)[10], Balk (2024)[8] and Diewert and Shimizu (2024)[31].

period. For comparison purposes, we also compute a simple index that uses the arithmetic average of prices in each region-month as an estimate of the price level for given month in the given region. We also compute unit value price indexes for comparison purposes.

In Section 4, we compute weighted and unweighted Time Product Dummy (TPD) Hedonic price indexes. In Section 5, we compute Geary-Khamis (GK). All three of these indexes are consistent with purchasers having linear preferences over rice products.

In Sections 6 and 7, we turn to the econometric estimation of linear preferences using share equations. In Section 6, we indicate why the estimation of the dual unit cost function is not practical when there are missing products for some region-months.

In Section 8, we estimate a Constant Elasticity of Substitution (CES) utility function using share equations and we compute the resulting price and quantity indexes for the 6 Prefectures. The estimation process also gives us an estimate for the elasticity of substitution and we use this estimate to implement Feenstra's (1994)[35] method for measuring the welfare benefits and costs of differing choice sets across the Prefectures. We find that these welfare effects are substantial. However, all of our Prefecture price indexes show that variations in price levels across the 6 Prefectures are even more substantial; i.e., the lower population Prefectures tend to have much higher price levels for rice than the corresponding Tokyo levels. This has implications for constructing national price and volume estimates for the System of National Accounts; i.e., the usual national consumption price deflators may be biased (because they do not take into account the fact that price levels may differ substantially by region and hence are biased).

Finally, in Section 9, we estimate a more flexible functional form for the purchasers utility function, the KBF (Konüs-Byushgens-Fisher) utility function with a rank 1 substitution matrix. The CES functional form is of course more flexible than the linear utility function but it has only a single parameter (the elasticity of substitution) to describe substitution possibilities. If there are N products that are in scope, the KBF functional form has $N-1$ parameters to describe substitution possibilities so it is much more flexible than the CES functional form.

We will explain each of the above methods in the following sections.

Section 10 concludes.

2 The Japanese Data and Notation

The six Prefectures are Hokkaido, Tokyo, Kyoto, Tottori, Kochi and Kagoshima. Over the 6 regions and 24 months, there were 499 separate rice products purchased in shops[*9] over this sample period. However, we excluded small sales of products that had very low sales so that we considered only 80 best selling rice dish products. The percentage of total purchase value by the 80 top selling products was 88%. There was a maximum total of $6 \times 24 \times 80 = 11,520$ observations of monthly product sales. However, many products were missing in each region so the probability that any one of the 80 products was available in any of the 144 region-months was equal to 0.5642.[*10]

[*9] We use the weekly retail sales database, "SRI+®(Nationwide Retail Store Panel Survey)" by INTAGE Inc. The analysis covered the period from the first week of January 2021 to the last week of December 2022, and included 804 supermarkets in Hokkaido, Tokyo, Kyoto, Tottori, Kochi, and Kagoshima that sold products in the Rice Dish category. For the data used in the analysis, first, the data for each store was aggregated by prefecture, and then the weekly data was aggregated into monthly data to create monthly data on sales amounts and sales quantities by product and prefecture.

[*10] Here are the probabilities of a product being present in each of the 6 Prefectures: 0.646 0.898 0.730 0.239 0.408 0.465. Thus for Tottori, the probability any one of the 80 products being purchased in any of the

Here is a description of the notation we will use for prices, quantities and values for our data. Denote the unit value price and total quantity purchased of product n in region r and month m by p_{rmn} and q_{rmn} for $r = 1, \ldots, 6; m = 1, 2, \ldots, 24$ and $n = 1, \ldots, 80$. If product n in region r and month m was not purchased, then we set $p_{rmn} = 0$ and $q_{rmn} = 0$.

Define the price and quantity vectors for region r and month m as $\boldsymbol{p}^{r,m} \equiv [p_{rm1}, \ldots, p_{rm80}]$ and $\boldsymbol{q}^{r,m} \equiv [q_{rm1}, \ldots, q_{rm80}]$ for $r = 1, \ldots, 6$ and $m = 1, \ldots, 24$.

Relabel the 144 region-month price vectors of dimension 80, $[\boldsymbol{p}^{1,1}, \ldots, \boldsymbol{p}^{1,24}; \ldots; \boldsymbol{p}^{6,1}, \ldots, \boldsymbol{p}^{6,24}]$, as $[\boldsymbol{p}^{1}, \ldots, \boldsymbol{p}^{24}; \ldots; \boldsymbol{p}^{121}, \ldots, \boldsymbol{p}^{144}]$ and relabel the 144 region-month quantity vectors $[\boldsymbol{q}^{1,1}, \ldots, \boldsymbol{q}^{1,24}; \ldots; \boldsymbol{q}^{6,1}, \ldots, \boldsymbol{q}^{6,24}]$ as $[\boldsymbol{q}^{1}, \ldots, \boldsymbol{q}^{24}; \ldots; \boldsymbol{q}^{121}, \ldots, \boldsymbol{q}^{144}]$. We also relabel the nth component of $\boldsymbol{p}^t$ as p_{tn} and the nth component of $\boldsymbol{q}^t$ as q_{tn} for $n = 1, \ldots, 80$.

Thus we have relabelled the 144 region-month price and quantity vectors , $\boldsymbol{p}^{r,m}$ and $\boldsymbol{q}^{r,m}$, into 144 time period vectors, $\boldsymbol{p}^t$ and $\boldsymbol{q}^t$, for $t = 1, \ldots, 144$. This relabelling simplified the computer programs. The corresponding expenditure or value vector for period t, $\boldsymbol{v}^t$, was obtained by multiplying together the components of $\boldsymbol{p}^t$ with the corresponding components of $\boldsymbol{q}^t$.

The first 24 period price vectors of dimension 80, $\boldsymbol{p}^t$ for $t = 1, \ldots, 24$, correspond to the month 1 to month 24 price vectors for region 1 (Hokkaido) and the last 24 period price vectors $\boldsymbol{p}^t$, $t = 121, \ldots, 144$, correspond to the month 1 to 24 price vectors for region 6 (Kagoshima).

Define the set of purchased products n in period t as $S(t)$ for $t = 1, \ldots, 144$.

3 Bilateral Fisher Indexes and GEKS Multilateral Indexes

The multilateral GEKS method is due to Gini (1924)[42] (1931)[43] and was further developed by Eltetö and Köves (1964)[34] and Szulc (1964)[64].

In order to define the GEKS indexes, we first need to define the Laspeyres, Paasche and Fisher (1922)[36] bilateral price indexes. These indexes which compare the prices of period t to the prices of period s, $P_{\text{L}}(t/s)$, $P_{\text{P}}(t/s)$ and $P_{\text{F}}(t/s)$ are defined as follows (using our new notation explained above):

$$P_{\text{L}}(t/s) \equiv \textstyle\sum_{n \in S(t) \cap S(s)} p_{tn} q_{sn} / \sum_{n \in S(t) \cap S(s)} p_{sn} q_{sn}; \qquad 1 \leq s, t \leq 144; \tag{1}$$

$$P_{\text{P}}(t/s) \equiv \textstyle\sum_{n \in S(t) \cap S(s)} p_{tn} q_{tn} / \sum_{n \in S(t) \cap S(s)} p_{sn} q_{tn}; \qquad 1 \leq s, t \leq 144; \tag{2}$$

$$P_{\text{F}}(t/s) \equiv [P_{\text{L}}(t/s) P_{\text{P}}(t/s)]^{1/2}; \qquad 1 \leq s, t \leq 144. \tag{3}$$

Note that the comparison of prices in period t to the prices in period s is restricted to products that were purchased in both periods; i.e., the above bilateral indexes are based on *matched product prices.*

The sequence of Fisher price indexes that compare the prices of period $1, 2, \ldots, 144$ with the prices of period s is $[P_{\text{F}}(1/s), P_{\text{F}}(2/s), \ldots, P_{\text{F}}(144/s)]$. This sequence of numbers is the *Fisher star index* for the 144 periods relative to the fixed base of period s. Obviously, any period could be chosen as the base so there are 144 Fisher star indexes. The GEKS index simply takes the geometric mean of these indexes. Thus the (preliminary) GEKS price level for period t, P^t_{GEKSP}, is defined as follows:

$$P^t_{\text{GEKSP}} \equiv [\textstyle\prod_{s=1}^{144} P_{\text{F}}(t/s)]^{1/144}; \qquad t = 1, \ldots, 144. \tag{4}$$

24 months was only 23.9% while for Tokyo, the corresponding probability was 89.8%. This is a huge variation in product availability. The range of products available over the 24 months in each region was: Region 1: 48-53; Region 2: 70-74; Region 3; 54-62; Region 4: 18-20; Region 5: 28-37; Region 6: 36-38. It can be seen that the lack of matching prices across Prefectures is a big problem.

We normalize the above 144 price levels so that the price level for period 25 (equal to the price level for Tokyo in month 1 of our sample of 24 months) is set equal to 1. Thus the final GEKS price level for period t is:

$$P^t_{\mathrm{GEKS}} \equiv P^t_{\mathrm{GEKSP}}/P^{25}_{\mathrm{GEKSP}}; \qquad t = 1, \ldots, 144. \tag{5}$$

Finally, we decompose the P^t_{GEKS} vector of dimension 144 into 6 regional vectors of dimension 24: $P^1_{\mathrm{GEKS}}, P^2_{\mathrm{GEKS}}, \ldots, P^6_{\mathrm{GEKS}}$. Component t of these GEKS price index vectors for regions 1 to 6 are defined as follows:

$$P^{1,m}_{\mathrm{GEKS}} \equiv P^m_{\mathrm{GEKS}}; P^{2,m}_{\mathrm{GEKS}} \equiv P^{24+m}_{\mathrm{GEKS}}; \ldots; P^{6,m}_{\mathrm{GEKS}} \equiv P^{120+m}_{\mathrm{GEKS}}; \quad m = 1, \ldots, 24. \tag{6}$$

The Fisher star index that uses month 1 for Tokyo as the base month that has the 144 components $P_{\mathrm{F}}(t/25)$ for $t = 1, \ldots, 144$ is also decomposed into 6 regional vectors of dimension 24, which we denote by $P^1_{\mathrm{F}}, P^2_{\mathrm{F}}, \ldots, P^6_{\mathrm{F}}$. Thus the 24 components of these 6 vectors are defined as follows:

$$P^{1,m}_{\mathrm{F}} \equiv P_{\mathrm{F}}(m/25); P^{2,m}_{\mathrm{F}} \equiv P_{\mathrm{F}}((24+m)/25); \ldots; P^{6,m}_{\mathrm{F}} \equiv P_{\mathrm{F}}((120+m)/25); \quad m = 1, \ldots, 24. \tag{7}$$

The Fixed Base Fisher Star indexes with month 1 for Tokyo as the base and the GEKS indexes for region r, P^r_{F} and P^r_{GEKS} for $r = 1, \ldots, 6$ are listed in Table 1 in Appendix and plotted on Chart 1 below as P^t_{GEKS} for $t = 1, \ldots, 144$. On Chart 1, we stacked the 6 Prefecture Fixed Base Fisher price indexes P^r_{F} of dimension 24 into a single index P^t_{F} for $t = 1, \ldots, 144$. We also stacked the other 4 indexes described in this section into single indexes of dimension 144 in order to reduce the number of charts.

A possible problem with the fixed base Fisher indexes (and with the GEKS indexes) is that over time, product matches may decrease due to product churn and hence the indexes become less reliable. A possible solution to this lack of matching problem could be to construct chained Fisher indexes for the most populous region, Tokyo, and then at each month, multiply the chained Tokyo index by the fixed base Fisher index linking the regional prices for Prefectures 1, 3, 4, 5 and 6 to the Tokyo prices in the same month.[*11] We explain the algebra for this method in the following paragraph.

First, it is necessary to construct the chained Fisher indexes for Tokyo, P^m_{FCH} for months $m = 1, \ldots, 24$. For month 1, define $P^1_{\mathrm{FCH}} = 1 = P_{\mathrm{F}}(25/25)$. For month 2, define $P^2_{\mathrm{FCH}} = P^1_{\mathrm{FCH}} \times P_{\mathrm{F}}(26/25)$; for month 3, define $P^3_{\mathrm{FCH}} = P^2_{\mathrm{FCH}} \times P_{\mathrm{F}}(27/26)$; ... : for month 24, define $P^{24}_{\mathrm{FCH}} = P^{23}_{\mathrm{FCH}} \times P_{\mathrm{F}}(48/47)$. The Fisher-Misobuchi multilateral indexes for month m for Prefecture 2 are defined as $P^{2,m}_{\mathrm{FM}} \equiv P^m_{\mathrm{FCH}}$ for $m = 1, \ldots, 24$. Thus the FM indexes for Tokyo are simply the chained Fisher indexes P^m_{FCH}. The Fisher-Misobuchi indexes for month m for Prefecture 1 are defined as $P^{1,m}_{\mathrm{FM}} \equiv P^m_{\mathrm{FCH}} \times P_{\mathrm{F}}(m/(24+m))$ for $m = 1, \ldots, 24$. The Fisher-Misobuchi indexes for month m for Prefecture 3, 4, 5 and 6 are defined as $P^{3,m}_{\mathrm{FM}} \equiv P^m_{\mathrm{FCH}} \times P_{\mathrm{F}}((48+m)/(24+m))$, $P^{4,m}_{\mathrm{FM}} \equiv P^m_{\mathrm{FCH}} \times P_{\mathrm{F}}((72+m)/(24+m))$, $P^{5,m}_{\mathrm{FM}} \equiv P^m_{\mathrm{FCH}} \times P_{\mathrm{F}}((96+m)/(24+m))$ and $P^{6,m}_{\mathrm{FM}} \equiv P^m_{\mathrm{FCH}} \times P_{\mathrm{F}}((120+m)/(24+m))$ for $m = 1, \ldots, 24$. Denote the vectors of dimension 24 of these multilateral indexes for Prefectures 1-6 by $P^1_{\mathrm{FM}}, P^2_{\mathrm{FM}}, P^3_{\mathrm{FM}}, P^4_{\mathrm{FM}}, P^5_{\mathrm{FM}}, P^6_{\mathrm{FM}}$. These indexes are listed in Appendix and stacked and plotted as P^t_{FM} for $t = 1, \ldots, 144$ on Chart 1 below.

It is also of interest to construct unit value prices and average prices for each region-month in our sample. Let $\mathbf{1}_{80}$ be a vector of ones of dimension 80. The unit value price for period t

[*11] This multilateral method was suggested by Hideyuki Mizobuchi.

is equal to period t expenditure on the 80 products, $\boldsymbol{v}^t \cdot \mathbf{1}_{80}$, divided by the number of units of rice purchased in period t, $\boldsymbol{q}^t \cdot \mathbf{1}_{80}$, where $\boldsymbol{v}^t \cdot \mathbf{1}_{80} \equiv \sum_{n \in S(t)} v_{tn}$ and $\boldsymbol{q}^t \cdot \mathbf{1}_{80} \equiv \sum_{n \in S(t)} q_{tn}$ for $t = 1, \ldots, 144$. We divide these 144 period unit value prices by the period 25 unit value price level to make up our unit value price index for all 144 periods, with month 1 for Tokyo as our base period where the unit value index is equal to 1. We then decompose this vector of dimension 144 into 6 vectors of dimension 24 to define the 6 regional unit value price indexes, $P^1_{\mathrm{UV}}, P^2_{\mathrm{UV}}, \ldots, P^6_{\mathrm{UV}}$. These unit value price indexes for the 6 regions have the 24 components $P^{r,m}_{\mathrm{UV}}$ defined as follows:[*12]

$$
\begin{aligned}
P^{1,m}_{\mathrm{UV}} &\equiv [\boldsymbol{v}^m \cdot \mathbf{1}_{80}/\boldsymbol{q}^m \cdot \mathbf{1}_{80}]/[\boldsymbol{v}^{25} \cdot \mathbf{1}_{80}/\boldsymbol{q}^{25} \cdot \mathbf{1}_{80}]; \qquad m = 1, \ldots, 24 \\
P^{2,m}_{\mathrm{UV}} &\equiv [\boldsymbol{v}^{(m+24)} \cdot \mathbf{1}_{80}/\boldsymbol{q}^{(m+24)} \cdot \mathbf{1}_{80}]/[\boldsymbol{v}^{25} \cdot \mathbf{1}_{80}/\boldsymbol{q}^{25} \cdot \mathbf{1}_{80}]; \\
&\cdots \\
P^{6,m}_{\mathrm{UV}} &\equiv [\boldsymbol{v}^{(m+120)} \cdot \mathbf{1}_{80}/\boldsymbol{q}^{(m+120)} \cdot \mathbf{1}_{80}]/[\boldsymbol{v}^{25} \cdot \mathbf{1}_{80}/\boldsymbol{q}^{25} \cdot \mathbf{1}_{80}]. \qquad (8)
\end{aligned}
$$

It is also of interest to construct a simple (unweighted) Average Price index for each region-period. The average price of a rice product purchased in period t is defined as the sum of the period t prices, , $\boldsymbol{p}^t \cdot \mathbf{1}_{80}$, divided by the number of rice products that were purchased in period t, N^t, for $t = 1, \ldots, 144$. We divide these 144 period average prices by the period 25 average price to make up our average price index for all 144 periods. We then decompose this vector of dimension 144 into 6 vectors of dimension 24 to define the 6 regional average price indexes, $P^1_{\mathrm{AV}}, P^2_{\mathrm{AV}}, \ldots, P^6_{\mathrm{AV}}$. These Average Price indexes for regions 1-6, have 24 components defined as follows:

$$
\begin{aligned}
P^{1,m}_{\mathrm{AV}} &\equiv [\boldsymbol{p}^m \cdot \mathbf{1}_{80}/N^m]/[\boldsymbol{p}^{25} \cdot \mathbf{1}_{80}/N^{25}]; \qquad m = 1, \ldots, 24 \\
P^{2,m}_{\mathrm{AV}} &\equiv [\boldsymbol{p}^{(m+24)} \cdot \mathbf{1}_{80}/N^{(m+24)}]/[\boldsymbol{p}^{25} \cdot \mathbf{1}_{80}/N^{25}]; \\
&\cdots \\
P^{6,m}_{\mathrm{AV}} &\equiv [\boldsymbol{p}^{(m+132)} \cdot \mathbf{1}_{80}/N^{(m+120)}]/[\boldsymbol{p}^{25} \cdot \mathbf{1}_{80}/N^{25}]. \qquad (9)
\end{aligned}
$$

These indexes are listed in Appendix and stacked into the index P^t_{AV} for $t = 1, \ldots, 144$ and plotted on Chart 1 below.

It can be seen the 5 Tokyo indexes (observations $t =$25-48) are all fairly close. The Average Price index levels, P^t_{AV}, tend to be smoother than the other indexes but the P^t_{AV} levels for Tottori (observations $t =$73-96) are far below the other indexes. The Unit Value indexes, P^t_{UV}, are very volatile and for the most part, lie below the other indexes. We will not consider P^t_{AV} and P^t_{UV} in the remainder of the paper. It can be seen that price levels tended to be a bit lower than Tokyo price levels for Regions 1 and 3 (Hokkaido and Kyoto) but price levels tended to be higher in the smaller population Regions 4-6 (Tottori, Kochi and Kagoshima) and particularly high in Tottori. The three indexes that utilize bilateral Fisher indexes in their construction are somewhat close to each other but there is a great deal of volatility in these indexes, particularly for the low population Prefectures.

The GEKS price levels are transitive (or path independent) and are invariant to changes in the units of measurement. However, they do not satisfy a strong version of Walsh's (1901; 389)[67] (1921; 540)[68] Multiperiod Identity Test which asks that the price levels for any two periods

[*12] Notation: if $\boldsymbol{a} \equiv [a_1, a_2, \ldots, a_N]$ and $\boldsymbol{b} \equiv [b_1, b_2, \ldots, b_N]$ are vectors of dimension N, then $\boldsymbol{a} \cdot \boldsymbol{b} \equiv \sum_{n=1}^{N} a_n b_n$.

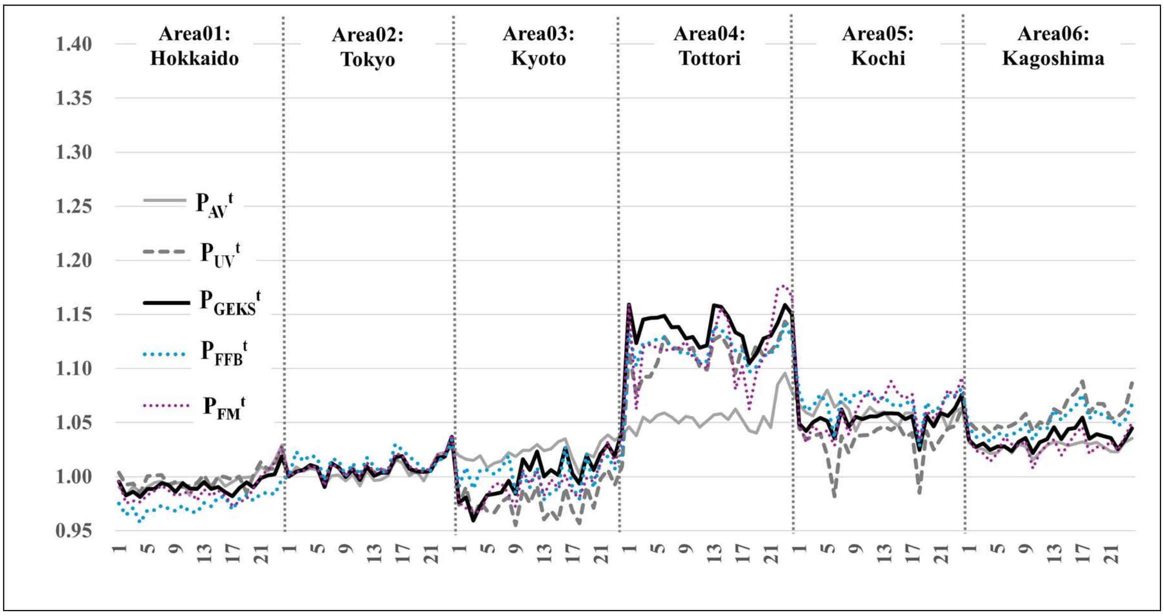

Chart 1 Average Price, Unit Value, GEKS, Fixed Base Fisher and Mizobuchi Fisher Price Indexes for Six Japanese Prefectures

be equal if the underlying price vectors are identical for the two periods.[*13] When there are missing prices, these GEKS indexes that are based on matched products are no longer exact for a flexible functional form.[*14]

Here are some tentative conclusions that emerge from the above Chart:

- Price levels were somewhat stable on average for each Prefecture until the last 5 months of 2023 when prices rose quite rapidly.
- Price levels differed substantially across Prefectures; the higher population Prefectures 1-3 had similar fairly stable price levels for the first 20 months in our sample and then experienced rapid inflation. The smaller population Prefectures 4-6 had substantially higher price levels compared to the Tokyo levels throughout the sample period.

We will exclude the Average Price and Unit Value Price indexes, P_{AV}^t and P_{UV}^t, from further

[*13] Suppose that there are no missing products and the price and quantity vectors for 3 periods are $\boldsymbol{p}^t, \boldsymbol{q}^t$ for $t = 1, 2, 3$. Suppose that the price vectors for periods 1 and 2 are identical so we have $\boldsymbol{p}^1 = \boldsymbol{p}^2 \equiv \boldsymbol{p}$. Then using definitions (4) adapted to the current context and straightforward calculations show that the ratio of the GEKS price levels for periods 1 and 2 is $P_{\text{GEKS}}^1/P_{\text{GEKS}}^2 = \{[\boldsymbol{p}\cdot\boldsymbol{q}^1/\boldsymbol{p}\cdot\boldsymbol{q}^2]^2[\boldsymbol{p}^3\cdot\boldsymbol{q}^2/\boldsymbol{p}^3\cdot\boldsymbol{q}^1]\}^{1/6}$, which is not equal to 1 in general. However, this ratio *is* equal to 1 if $\boldsymbol{q}^1 = \boldsymbol{q}^2$. Thus the GEKS price levels for periods 1 and 2 are identical if prices *and* quantities are equal for periods 1 and 2, so the GEKS price levels satisfy a *weak identity test*. If we interchange prices and quantities in the above algebra, we can show that the GEKS quantity levels satisfy a weak identity test but not the strong identity test. For more on the Test Approach to Multilateral Price or Quantity Levels, see Zhang, Johansen and Nygaard (2019)[70] and Diewert (2023)[25].

[*14] See Diewert (1976)[18] for materials on exactness of index numbers for flexible functional forms. The results in Diewert assumed that there were no missing prices or unavailable products in any two periods that were being compared. In the present situation, there are a great many missing prices. If we had reservation prices for the products that were unavailable, then Diewert's results would be valid. But we do not have reservation prices in hand so bilateral Fisher indexes are computed only over products that are present in the two periods that are being compared. The resulting GEKS indexes (that are constructed using bilateral comparisons) are no longer exact for a flexible functional form.

consideration as "best" indexes due to the unrepresentative nature of P^t_{AV} and the volatility of P^t_{UV}. The remaining indexes do not satisfy strong identity tests.*15

4 Weighted and Unweighted Time Product Dummy Hedonic Regressions

We have converted the 144 region-month price, quantity and value vectors into 144 $\boldsymbol{p}^t, \boldsymbol{q}^t$ and $\boldsymbol{v}^t$ vectors of dimension 80 that are indexed by an artificial time index t for $t = 1, \ldots, 144$. Thus we can run a weighted or unweighted Time-Product-Dummy hedonic regression to estimate the 144 rice price levels for the 144 region-months in our sample.

The unweighted TDP model dates back to Court (1939)[11] and Summers (1973)[63]; the Weighted TPD model dates back to Rao (1995)[60] (2005)[61] and Diewert (2004)[20] (2005)[21].

These models are based on the price data satisfying (to some degree of approximation) the following equations:

$$p_{tn} \approx \pi_t \alpha_n; \qquad t = 1, \ldots, 144; n \in S(t) \tag{10}$$

where π_t is interpreted as the period t price level and α_n is a parameter which reflects the quality (or marginal utility) of product n. Thus π_t is a summary measure for the level of prices in the region-month that corresponds to period t.

Taking logarithms of both sides of the approximate equalities defined by (10) leads to the following approximate equalities:

$$\begin{aligned} \ln p_{tn} &\approx \ln \pi_t + \ln \alpha_n; \qquad t = 1, \ldots, 144; n \in S(t) \\ &= \rho_t + \beta_n \end{aligned} \tag{11}$$

where $\rho_t \equiv \ln \pi_t$ for $t = 1, \ldots, 144$ and $\beta_n \equiv \ln \alpha_n$ for $n = 1, \ldots, 80$. The second set of approximate equations in (11) is a linear regression model. However, the ρ_t and β_n parameters in (11) are not uniquely determined; we require a normalization on one of these parameters. Choose the following normalization:

$$\rho_1 = 0 \ \text{ (which corresponds to } \pi_1 = 1). \tag{12}$$

The linear regression that corresponds to (11) and (12) is equivalent to solving the following least squares minimization problem (using also the normalization (12)):

$$\min_{\rho' s\, \beta' s} \{ \textstyle\sum_{t=1}^{144} \sum_{n \in S(t)} [\ln p_{tn} - \rho_t - \beta_n]^2 \}. \tag{13}$$

Assume for the moment that the approximate equations (11) hold as exact equations. Multiply both sides of equation p_{tn} by q_{tn} for $n \in S(t)$ and sum the resulting equations over n. For each t, we obtain the following equation:

$$\textstyle\sum_{n \in S(t)} p_{tn} q_{tn} = \boldsymbol{p}^t \cdot \boldsymbol{q}^t \equiv e^t = \pi_t \sum_{n \in S(t)} \alpha_n q_{tn} = \pi_t \boldsymbol{\alpha} \cdot \boldsymbol{q}^t = \pi_t Q^t; \quad t = 1, \ldots, 144; \tag{14}$$

where $e^t \equiv \sum_{n \in S(t)} p_{tn} q_{tn} = \sum_{n=1}^{80} p_{tn} q_{tn} \equiv \boldsymbol{p}^t \cdot \boldsymbol{q}^t$ is defined to be period t expenditure on the 80 products and the period t quantity aggregate $Q^t \equiv \boldsymbol{\alpha} \cdot \boldsymbol{q}^t \equiv \sum_{n \in S(t)} \alpha_n q_{tn}$ for $t = 1, \ldots, 144$.

*15 It is of interest to compare the GEKS indexes to the indexes that will be defined in subsequent sections.

Thus if equations (10) or (11) hold as equalities rather than as approximate equalities, the Time Product Dummy Hedonic Regression model implicitly assumes that purchasers of the rice products have the same linear preferences $f(\boldsymbol{q}) \equiv \boldsymbol{\alpha} \cdot \boldsymbol{q}$ over the 80 products.*16

Of course, equations (11) and (12) will not hold as exact equalities. For our data set, we ran the linear regression defined by equations (11) and the normalization (12). We obtained estimates for the 143 nonzero ρ_t parameters which we denote as ρ_t^* for $t = 2, 3, \ldots, 144$. We defined $\rho_1^* \equiv 0$.

The estimated price levels for our 144 region-months were defined as follows:

$$\pi_t^* \equiv \exp[\rho_t^*]; \qquad t = 1, \ldots, 144. \tag{15}$$

The corresponding period t quantity levels Q^{t^*} were defined by deflating period t expenditures e^t by the π_t^*:

$$Q^{t^*} \equiv e^t/\pi_t^*; \qquad t = 1, \ldots, 144. \tag{16}$$

We used definitions (15) to define the regions 1-6 Time Product Dummy price levels P_{TPD}^{rt} for $r = 1, \ldots, 6$ and $t = 1, \ldots, 24$ as follows:

$$P_{\text{TPD}}^{1,m} \equiv \pi_t^*; P_{\text{TPD}}^{2,m} \equiv \pi_{t+24}^*; \ldots; P_{\text{TPD}}^{6,m} \equiv \pi_{t+120}^*; \qquad m = 1, \ldots, 24. \tag{17}$$

These indexes are listed in Table 2 of the Appendix. They are stacked into the single index P_{TPD}^t of dimension 144 which is plotted on Chart 2 below. The stacked GEKS price index P_{GEKS}^t is also plotted on Chart 2 for comparison purposes.

If expenditure information is available, then it is possible to use the above hedonic regression results to decompose period t expenditure e^t into price and quantity components in an alternative way. For products n that were purchased in at least one region-month, we have nonzero estimates for the logarithms of the quality adjustment parameters, which we denote by β_n^*. For these products, define the corresponding quality adjustment parameter $\alpha_n^* \equiv \exp[\beta_n^*]$. Define $\boldsymbol{\alpha}^*$ as the resulting vector of the α_n^*, $\boldsymbol{\alpha}^* \equiv [\alpha_1^*, \ldots, \alpha_{80}^*]$. Define the period t alternative price and quantity levels, $P^{t^{**}}$ and $Q^{t^{**}}$ as follows:

$$Q^{t^{**}} \equiv \boldsymbol{\alpha}^* \cdot \boldsymbol{q}^t; P^{t^{**}} \equiv e^t/Q^{t^{**}}; \qquad t = 1, \ldots, 144. \tag{18}$$

If the fit in the regressions defined by (10) or (11) is perfect, then the two sets of aggregate price and quantity levels will coincide.*17

An advantage of the unweighted (or more properly, the equally weighted) TPD price indexes defined by (15) is that they can be constructed using just price information on the purchased

---

*16 The estimating equations (10) for the parameters π_t and α_n could be written as $p_{tn} = \pi_t \alpha_n + \varepsilon_{tn}$ for $t = 1, \ldots, 144$ and $n \in S(t)$ where the ε_{tn} are error terms. Then the following equations are the counterparts to equations (14) that take the error terms into account: $e^t = \pi_t(\sum_{n \in S(t)} \alpha_n q_{tn}) + \varepsilon^t = \pi_t \boldsymbol{\alpha} \cdot \boldsymbol{q}^t + \varepsilon^t$ where $\varepsilon^t \equiv (\sum_{n \in S(t)} \varepsilon_{tn} q_{tn})$ for $t = 1, \ldots, 144$. We want to decompose period t expenditure e^t into $P^t Q^t$ where P^t represents the period t aggregate price level and Q^t represents the aggregate quantity level. It is natural that we set $P^t \equiv \pi_t$ and $Q^t \equiv \boldsymbol{\alpha} \cdot \boldsymbol{q}^t$ but then the resulting product, $P^t Q^t$ will not equal actual expenditure on the products, e^t, if $\varepsilon^t \neq 0$. Thus if we choose $P^t \equiv \pi_t$, then choose the companion $Q^t = e^t/P^t$. If we choose $Q^t \equiv \boldsymbol{\alpha} \cdot \boldsymbol{q}^t$, then we choose the companion $P^t = e^t/Q^t$. This ensures that no value disappears or is created by the aggregation process; i.e., P^t and Q^t satisfy the Product Test for levels: $P^t Q^t = e^t$.

*17 Use the estimates α_n^* and π_t^* to define the error terms $e_{tn} \equiv p_{tn} - \pi_t^* \alpha_n^*$ for $t = 1, \ldots, 144$; $n \in S(t)$. If the period t error terms e_{tn} sum to zero, so that $\sum_{n \in S(t)} e_{tn} = 0$, then it can be shown that $P^{t^{**}} = P^{t^*}$ and $Q^{t^{**}} = Q^{t^*}$.

rice products. But this is also a disadvantage of these estimates because products with low volumes of sales should not get the same weight in the regression as highly popular products. For discussions on the benefits and costs of alternative weighting methods, see Diewert (2004)[20] (2023)[25].

The weighting method that we considered in this paper is defined by the following Weighted Time Product Dummy Least Squares estimation model that is a weighted counterpart to the TPD model defined by the least squares minimization problem defined by (13) above:

$$\min_{\rho' s\, \beta' s}\{\textstyle\sum_{t=1}^{144}\sum_{n\in S(t)} s_{tn}[\ln p_{tn} - \rho_t - \beta_n]^2\} \tag{19}$$

where $s_{tn} \equiv p_{tn}q_{tn}/\boldsymbol{p}^t \cdot \boldsymbol{q}^t$ is the expenditure share of product n in period t.

We still require a normalization like $\rho_1 = 0$ to get a unique solution to the least squares minimization problem defined by (19).

To compute the solution to (19), we used the following weighted version of the linear regression equations (11):

$$[s_{tn}]^{1/2} \ln p_{tn} \approx [s_{tn}]^{1/2}[\rho_t + \beta_n]; \qquad t = 1, \ldots, 144; n \in S(t). \tag{20}$$

Once the linear regression defined by (20) was run, we exponentiated the estimated ρ_t^* to define the $\pi_t^* \equiv \exp[\rho_t^*]$ and then used definitions (15) and (18) to define the Weighted Time Dummy Product price levels P_{WTPD}^t for $t = 1, \ldots, 144$ and to define the regional price indexes, $P_{\mathrm{WTPD}}^{1,m}$ - $P_{\mathrm{WTPD}}^{6,m}$ for $m = 1, \ldots, 24$. These indexes are listed in Table 2 in the Appendix and the stacked indexes P_{WTPD}^t for $t = 1, \ldots, 144$ are plotted on Chart 2 below.

The Weighted Time Product Dummy price levels P_{WTPD}^t were defined *directly* by using the exponentials of the estimates for the price levels, the π_t^*, and then the companion period t quantity levels were defined *implicitly* by definitions (16), $Q^{t^*} \equiv e^t/\pi_t^*$ for $t = 1, \ldots, 144$. But as indicated above, it is possible to use the exponentials of the estimated quality adjustment parameters, the β_n^* for $n = 1, \ldots, N$. For these products, define the corresponding quality adjustment parameter $\alpha_n^* \equiv \exp[\beta_n^*]$. Define $\boldsymbol{\alpha}^*$ as the resulting vector of the α_n^*, $\boldsymbol{\alpha}^* \equiv [\alpha_1^*, \ldots, \alpha_{80}^*]$. Now use definitions (18) to define the period t quantity levels *directly* as $Q^{t^{**}} \equiv \boldsymbol{\alpha}^* \cdot \boldsymbol{q}^t$ and the corresponding period t price levels *indirectly* as $P^{t^{**}} \equiv e^t/Q^{t^{**}}$ for $t = 1, \ldots, 144$.[*18] Again, if the fit in the linear regression (20) is perfect, the direct and indirect estimates will coincide. The advantage of the indirect estimation procedure is that the resulting quantity levels, $Q^{t^{**}}$, will satisfy the *strong identity test*; i.e., if $\boldsymbol{q}^t = \boldsymbol{q}^\tau$, then $Q^{t^{**}} = Q^{\tau^{**}}$. The regional Implicit Weighted Time Product Dummy price levels P_{IWTPD}^{rm} are defined using the $P^{t^{**}}$ divided by $P^{25^{**}}$, the Tokyo Implicit price level for month 1:

$$P_{\mathrm{IWTPD}}^{1,m} \equiv P^{m^{**}}/P^{25^{**}}; P_{\mathrm{IWTPD}}^{2,m} \equiv P^{m+24^{**}}/P^{25^{**}}; \ldots; P_{\mathrm{IWTPD}}^{6,m} \equiv P^{m+120^{**}}/P^{25^{**}};$$
$$m = 1, \ldots, 24. \tag{21}$$

These regional indexes are listed in Table 2 in the Appendix and they are stacked into the index P_{IWTPD}^t which is plotted on Chart 2 below.[*19]

[*18] The two methods for constructing price and quantity levels from a hedonic regression was discussed by de Haan and Krsinich (2018; 766)[15] who showed that before normalization that the Implicit WTPD price levels are equal or less than the corresponding (Direct) WTPD price levels. These authors called a price level of the form $e^t/\boldsymbol{\alpha} \cdot \boldsymbol{q}^t$ a *quality adjusted unit value price level.*

[*19] Use the estimates α_n^* and π_t^* to define the error terms $e_{tn} \equiv p_{tn} - \pi_t^* \alpha_n^*$ for $t = 1, \ldots, 144; n \in S(t)$. If the period t error terms e_{tn} sum to zero, so that $\sum_{n\in S(t)} e_{tn} = 0$, then it can be shown that $P^{t^{**}} = P^{t^*}$ and $Q^{t^{**}} = Q^{t^*}$.

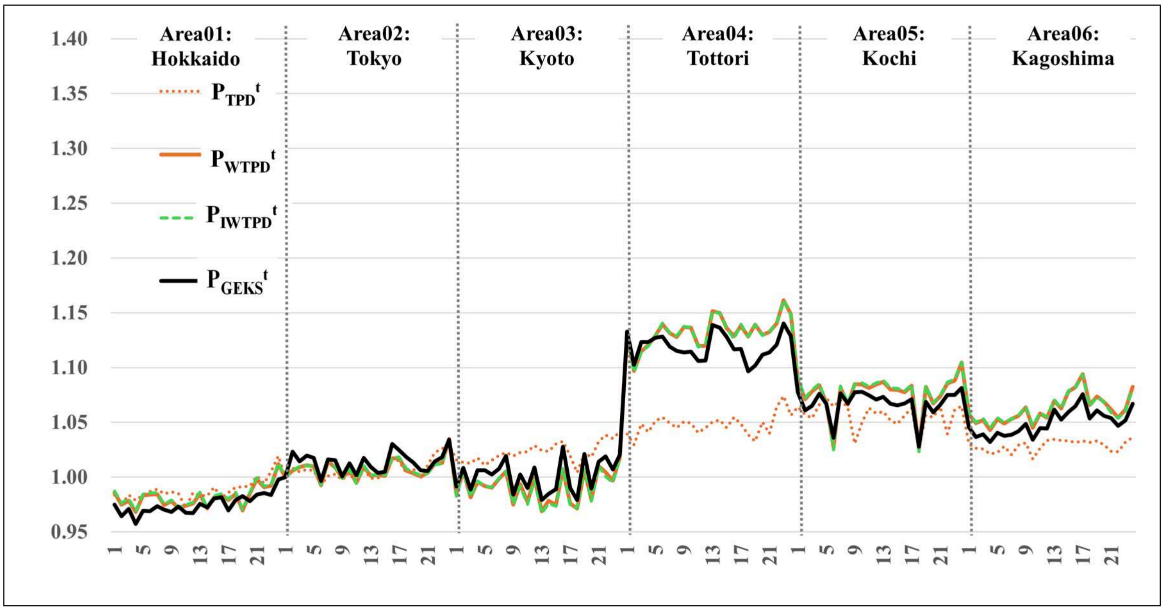

Chart 2 Unweighted, Weighted and Implicit Weighted Time Product Dummy Price Indexes

It can be seen that the (unweighted) Time Product Dummy indexes P^t_{TPD} are not close to the more appropriate weighted indexes, P^t_{WTPD}, P^t_{IWTPD} and P^t_{GEKS}, particularly for Prefectures 3-6 (observations 49-144). In particular, P^t_{TPD} for Tottori (observations 73-96) is far below the other 3 indexes. Indexes that are not weighted by economic importance can be unreliable.

The GEKS indexes, P^t_{GEKS}, are more volatile than the two Weighted TPD indexes, P^t_{WTPD} and P^t_{IWTPD}, which are very close and cannot be distinguished from each other on the Chart. In particular, the GEKS indexes are below the two Weighted Time Product Dummy indexes for the smaller Prefectures (observations 73-144).

In the following section, we introduce another multilateral index that is consistent with purchasers having linear preferences. This alternative index does not require econometric estimation.

5 Geary Khamis Multilateral Indexes

The GK multilateral method was introduced by Geary (1958)[41] in the context of making international comparisons of prices. Khamis (1970)[53] showed that the equations that define the method have a positive solution under certain conditions. A modification of this method has been adapted to the time series context and is being used to construct some components of the Dutch CPI; see Chessa (2016)[9]. The GK index was the multilateral index chosen by the Dutch to avoid the chain drift problem for the segments of their CPI that use scanner data.

Recall that $S(t)$ was the set of products n that were purchased in region-month t. Define $S^*(n)$ as the set of periods t where product n was sold. As was the case for the Time Product Dummy multilateral system of price and quantity levels, the equations which define the GK price and quantity levels involve 144 price levels π_t and 80 quality adjustment parameters α_n (recall equations (10) above). Define the vector $\boldsymbol{q}$ as the sum of the 144 observed quantity vectors $\boldsymbol{q}^t$ for each region-month t:

$$\boldsymbol{q} \equiv \sum_{t=1}^{144} \boldsymbol{q}^t. \tag{22}$$

The equations which determine the *GK price levels* $\pi_1, \ldots, \pi_{144}$ and *quality adjustment factors* $\alpha_1, ..., \alpha_{80}$ (up to a scalar multiple) are the following ones:

$$\alpha_n = \sum_{t \in S^*(n)} [q_{tn}/q_n][p_{tn}/\pi_t] = \sum_{n=1}^{80} [1/q_n][p_{tn}q_{tn}][1/\pi_t]; \qquad n = 1, ..., 80; \tag{23}$$

$$\pi_t = \boldsymbol{p}^t \cdot \boldsymbol{q}^t / \alpha \cdot \boldsymbol{q}^t = e^t / \alpha \cdot \boldsymbol{q}^t; \qquad t = 1, ..., 144 \tag{24}$$

where $\alpha \equiv [\alpha_1, ..., \alpha_N]$ is the vector of GK quality adjustment factors and $e^t \equiv \boldsymbol{p}^t \cdot \boldsymbol{q}^t$ is region-period t expenditure on the 80 rice products. Once a solution α and $\pi_1, \ldots, \pi_{144}$ to equations (23) and (24) has been found, the period t price levels P^t can be set equal to the corresponding π_t and the period t quantity levels are defined as follows:

$$Q^t \equiv \alpha \cdot \boldsymbol{q}^t; \qquad t = 1, \ldots, 144. \tag{25}$$

It can be seen that if a solution to equations (23) and (24) exists, then if all of the period price levels π_t are multiplied by a positive scalar λ say and all of the quality adjustment factors α_n are divided by the same λ, then another solution to (23) and (24) is obtained. Hence, the α_n and π_t are only determined up to a scalar multiple and an additional normalization is required such as $\pi_1 = 1$ or $\alpha_1 = 1$ is required to determine a unique solution to the system of equations defined by (23) and (24). The GK price and quantity levels have some good axiomatic properties including invariance to changes in the units of measurement.*20

A traditional method for obtaining a solution to (23) and (24) is to iterate between these equations. Thus set $\alpha = \mathbf{1}_{80}$, a vector of ones, and use equations (24) to obtain an initial sequence for the π_t. Substitute these π_t estimates into equations (23) and obtain α_n estimates. Substitute these α_n estimates into equations (24) and obtain a new sequence of π_t estimates. Continue iterating between the two systems until convergence is achieved.

Alternative methods are more efficient. Following Diewert (1999; 26)[19] and Diewert and Fox (2017; 31-32)[29], substitute equations (24) into equations (23) and after some simplification, obtain the following system of equations that will determine the components of the α vector (up to a scalar multiplicative factor):

$$[\mathbf{I}_N - \mathbf{C}]\alpha = \mathbf{0}_N \tag{26}$$

where $\mathbf{I}_N$ is the N by N identity matrix where $N = 80$, $\mathbf{0}_N$ is a vector of zeros of dimension N and the $\mathbf{C}$ matrix is defined as follows:

$$\mathbf{C} \equiv \hat{\boldsymbol{q}}^{-1} \sum_{t=1}^{T} \boldsymbol{s}^t \boldsymbol{q}^{tT} \tag{27}$$

where $\hat{\boldsymbol{q}}$ is an N by N diagonal matrix with the elements of the vector of total purchases $\boldsymbol{q}$ running down the main diagonal and $\hat{\boldsymbol{q}}^{-1}$ denotes the inverse of this matrix, $\boldsymbol{s}^t$ is the period t expenditure share column vector, $\boldsymbol{q}^t$ is the column vector of quantities purchased during period t and $\boldsymbol{q}^{tT}$ is the transpose of $\boldsymbol{q}^t$.

The matrix $\mathbf{I}_N - \mathbf{C}$ is singular which implies that the N equations in (26) are not all independent. In particular, if the first $N - 1$ equations in (26) are satisfied, then the last equation in (26) will also be satisfied. It can also be seen that the N equations in (26) are homogeneous of degree one in the components of the vector α. Thus to obtain a unique solution to (26),

*20 See Diewert (2023)[25] on the test properties of various multilateral indexes.

set α_N equal to 1, drop the last equation in (26) and solve the remaining $N-1$ equations for $\alpha_1, \alpha_2, ..., \alpha_{N-1}$. Once the α_n are known, equations (24) can be used to determine the GK price levels, $\pi_t = \boldsymbol{p}^t \cdot \boldsymbol{q}^t / \alpha \cdot \boldsymbol{q}^t$ for $t = 1, ..., 144$. This is the efficient procedure that was suggested and used by Diewert and Fox (2021)[30].

In the present study, we used another method to find a solution to equations (26) which also proved to be efficient. It can be seen that the matrix $\mathbf{C}$ has nonnegative elements. Hence under weak regularity conditions, $\mathbf{C}$ has a maximum positive eigenvalue (which can be shown to equal 1) and the associated eigenvector has strictly positive elements.*21 It was easy to form the matrix $\mathbf{C}$ using our data and Shazam's eigenvalue-eigenvector operator quickly generated the maximum eigenvalue (which of course equalled one) and the associated strictly positive eigenvector, which we denote by α^*. With α^* in hand, we calculated the 144 region-month quantity levels, $Q_{\text{GK}}^{t^*} \equiv \alpha^* \cdot \boldsymbol{q}^t$ and the associated price levels $P_{\text{GK}}^{t^*} = e^t / Q_{\text{GK}}^{t^*}$ for $t = 1, \ldots, 144$. These preliminary GK price levels $P_{\text{GK}}^{t^*}$ were divided by $P_{\text{GK}}^{25^*}$ (the preliminary GK price level for Tokyo in month 1) to obtain our 6 GK Prefecture price indexes, $P_{\text{GK}}^1, P_{\text{GK}}^2, \ldots, P_{\text{GK}}^6$. The components of these vectors of dimension 24, $P_{\text{GK}}^{r,m}$, are defined in two stages in the usual way:

$$P_{\text{GK}}^t \equiv P_{\text{GK}}^{t^*} / P_{\text{GK}}^{25^*}; \qquad t = 1, \ldots, 144; \tag{28}$$

$$P_{\text{GK}}^{1,m} \equiv P_{\text{GK}}^t; P_{\text{GK}}^{2,m} \equiv P_{\text{GK}}^{t+24}; \ldots; P_{\text{GK}}^{6,m} \equiv P_{\text{GK}}^{t+120}; \qquad t = 1, \ldots, 24. \tag{29}$$

The GK Prefecture price indexes, $P_{\text{GK}}^1, P_{\text{GK}}^2, \ldots, P_{\text{GK}}^6$, are listed in Table 3 in Appendix. They are stacked and plotted on Chart 3 in section 7 below.

We conclude this section by noting that the GK indexes defined by (23)-(25) are exact for linear preferences (products are perfect substitutes) and for Leontief preferences (products are not substitutable at all). The first result is obvious from definition (25), i.e., utility in period t, u^t, is defined to be equal to the aggregate quantity $Q^t \equiv \alpha \cdot \boldsymbol{q}^t$ for $t = 1, \ldots, T$ where T is equal to 144 in our empirical work. The second result was established by Diewert (1999; 58-60)[19] but his proof is quite complicated. It is possible to establish that the GK indexes are exactly consistent with all purchasers having Leontief preferences by using the simple proof below.

Consider the case where there are N products and T observations. Assume that the period t price and quantity vectors, $\boldsymbol{p}^t$ and $\boldsymbol{q}^t$ for $t = 1, \ldots, T$ are consistent with purchasers of the N products all having Leontief preferences. If we have Leontief preferences, then every product that is purchased in one period must be purchased in all periods.*22 This means that there exists an N dimensional vector of positive constants, $\boldsymbol{\beta}$, which has components $\beta_1, ..., \beta_N$, utility levels $Q^1, \ldots, Q^T$ and unit cost price levels π_t such that the following equations are satisfied:

$$\boldsymbol{q}^t = \boldsymbol{\beta} Q^t; \qquad t = 1, \ldots, T; \tag{30}$$

$$\pi_t = \boldsymbol{\beta} \cdot \boldsymbol{p}^t; \qquad t = 1, \ldots, T. \tag{31}$$

Denote $\mathbf{D}(\boldsymbol{q}^t)$ as the diagonal matrix with the elements of $\boldsymbol{q}^t$ on the main diagonal. With all products being positive in this case, equations (23)-(25) become the following equations:

$$\alpha = [\textstyle\sum_{t=1}^T \mathbf{D}(\boldsymbol{q}^t)]^{-1} [\sum_{t=1}^T \mathbf{D}(\boldsymbol{q}^t) \boldsymbol{p}^t / \pi_t]; \tag{32}$$

$$\pi_t = \boldsymbol{p}^t \cdot \boldsymbol{q}^t / \alpha \cdot \boldsymbol{q}^t; \qquad t = 1, \ldots, T; \tag{33}$$

$$Q^t \equiv \alpha \cdot \boldsymbol{q}^t; \qquad t = 1, \ldots, T. \tag{34}$$

*21 This result is due to Frobenius (1912)[39]; see Gantmacher (1959)[40] for a proof of this result and a discussion of the regularity conditions on $\mathbf{C}$ which ensure the result.

*22 Thus Leontief preferences are not consistent with missing prices or product churn.

Substitute equations (30) and (31) into equations (32) and we obtain the following vector equation:

$$
\begin{aligned}
\boldsymbol{\alpha} &= [\textstyle\sum_{t=1}^{T}\mathbf{D}(\boldsymbol{q}^t)]^{-1}[\textstyle\sum_{t=1}^{T}\mathbf{D}(\boldsymbol{q}^t)\boldsymbol{p}^t/\pi_t] \\
&= [\textstyle\sum_{t=1}^{T}\mathbf{D}(\boldsymbol{\beta})Q^t]^{-1}[\textstyle\sum_{t=1}^{T}\mathbf{D}(\boldsymbol{\beta})Q^t\boldsymbol{p}^t/\boldsymbol{\beta}\cdot\boldsymbol{p}^t] \\
&= [\textstyle\sum_{t=1}^{T}Q^t]^{-1}[\mathbf{D}(\boldsymbol{\beta})]^{-1}[\mathbf{D}(\boldsymbol{\beta})][\textstyle\sum_{t=1}^{T}Q^t\boldsymbol{p}^t/\boldsymbol{\beta}\cdot\boldsymbol{p}^t] \\
&= [\textstyle\sum_{t=1}^{T}Q^t]^{-1}[\textstyle\sum_{t=1}^{T}Q^t\boldsymbol{p}^t/\boldsymbol{\beta}\cdot\boldsymbol{p}^t]. \qquad (35)
\end{aligned}
$$

Thus the vector $\boldsymbol{\alpha}$ is well defined by (35), given that we know the variables that appear in (30) and (31). Take the inner product of both sides of equations (35) with $\boldsymbol{q}^r$ for $r = 1, \ldots, T$. Using equations (30), we obtain the following T equations:

$$
\begin{aligned}
\boldsymbol{\alpha}\cdot\boldsymbol{\beta}Q^r &= [\textstyle\sum_{t=1}^{T}Q^t]^{-1}[\textstyle\sum_{t=1}^{T}Q^t\boldsymbol{p}^t/\boldsymbol{\beta}\cdot\boldsymbol{p}^t]\cdot\boldsymbol{\beta}\,Q^r; \qquad r = 1, \ldots, T; \\
&= [\textstyle\sum_{t=1}^{T}Q^t]^{-1}[\textstyle\sum_{t=1}^{T}Q^t\boldsymbol{p}^t\cdot\boldsymbol{\beta}/\boldsymbol{\beta}\cdot\boldsymbol{p}^t]Q^r \\
&= [\textstyle\sum_{t=1}^{T}Q^t]^{-1}[\textstyle\sum_{t=1}^{T}Q^t]Q^r \\
&= Q^r. \qquad (36)
\end{aligned}
$$

Now normalize the α_n defined by (35) so that they satisfy the following constraint:

$$\boldsymbol{\alpha}\cdot\boldsymbol{\beta} = 1. \qquad (37)$$

The resulting GK indexes defined by (32)-(34) are exact for Leontief preferences.*23

In the following sections, we will turn our attention to indexes that are based on the econometric estimation of purchaser preferences. In the section that follows immediately, we explain why it is difficult to estimate dual representations of purchaser preferences when there are new and disappearing products.

6 The Estimation of Systems of Inverse Demand Functions

Traditional consumer demand theory in the case of homothetic or linearly homogeneous preferences works as follows:*24 assume a once differentiable functional form for the household unit cost function $c(\boldsymbol{p})$ (which is dual to the household linearly homogeneous utility function $f(\boldsymbol{q})$*25). Assume that in period t, all households have the same preferences and face the vector of period t prices $\boldsymbol{p}^t$. Suppose each household maximizes utility subject to a budget constraint. Let $\boldsymbol{q}^t$ be the observed vector of total purchases of the N products in scope and further assume that $\boldsymbol{q}^t$ is strictly positive. Let $e^t > 0$ be observed period t total expenditure on the products in scope. Then it can be shown that $\boldsymbol{q}^t, \boldsymbol{p}^t$ and e^t satisfy the following system of consumer demand functions:

$$\boldsymbol{q}^t = e^t\nabla c(\boldsymbol{p}^t)/c(\boldsymbol{p}^t); \qquad t = 1, \ldots, T \qquad (38)$$

*23 In our present context, Leontief preferences are not relevant since they imply that positive amounts of all products are purchased in all periods where the preferences of purchasers do not change.

*24 See Diewert (1974)[17]. The materials in this section are drawn from Diewert (2024)[26].

*25 We assume that $f(\boldsymbol{q})$ is a linearly homogeneous function so that the resulting price index is independent of the scale of the quantity vectors $\boldsymbol{q}^t$. We think that this is a reasonable assumption at the first stage of aggregation. It would be difficult for statistical agencies to produce price indexes that were conditional on the scale of purchaser demands.

where $\nabla c(\boldsymbol{p}^t)$ is the vector of first order partial derivatives of the unit cost function evaluated at $\boldsymbol{p}^t$.

However, if there are missing products in one or more periods in the sample period then there are problems with the above traditional consumer demand methodology. Suppose product n is not purchased in period t so that $q_{tn} = 0$. Then the nth component in equations (38) for period t becomes:

$$q_{tn} = 0 = e^t[\partial c(p_{t1}, \ldots, p_{tn}, \ldots, p_{tN})/\partial p_n]/c(p_{t1}, \ldots, p_{tn}, \ldots, p_{tN}). \quad (39)$$

The problem is that we cannot observe the price of product n in period t, p_{tn}. Conceptually, it is the Hicksian reservation price[*26] which is just high enough to deter households from purchasing the product. Thus for every missing product in the sample of periods, we need to estimate an unknown reservation price in order to apply traditional consumer demand theory. This is not workable in practice.[*27] Hausman (1996)[45] (1999)[46] used variants of this cost function methodology to estimate reservation prices but it is not known how he solved this estimation problem.

We turn to the estimation of the utility function, $f(\boldsymbol{q})$, instead of estimating the dual unit cost function. When we make this switch, it turns out that we get a "practical" system of estimating equations.

The inverse demand function estimation methodology starts with the assumption that the observed period t quantity vector $\boldsymbol{q}^t$ is a solution to the following period t utility maximization problem:

$$\max_q\{f(\boldsymbol{q}) : \boldsymbol{p}^t \cdot \boldsymbol{q} = e^t; \boldsymbol{q} \geq \boldsymbol{0}_N\}; \qquad t = 1, \ldots, T. \quad (40)$$

Assuming that the linearly homogeneous function f is differentiable, the first order conditions for the observed $\boldsymbol{q}^t$ to solve the period t purchaser utility maximization problem are the following conditions:

$$\nabla f(\boldsymbol{q}^t) = \lambda_t \boldsymbol{p}^t; \qquad t = 1, \ldots, T; \quad (41)$$

$$\boldsymbol{p}^t \cdot \boldsymbol{q}^t = e^t; \qquad t = 1, \ldots, T. \quad (42)$$

Take the inner product of both sides of (41) with $\boldsymbol{q}^t$ and solve the resulting equation for the Lagrange multiplier λ_t. We find that

$$\begin{aligned} \lambda_t &= \boldsymbol{q}^t \cdot \nabla f(\boldsymbol{q}^t)/e^t \qquad t = 1, \ldots, T \\ &= f(\boldsymbol{q}^t)/e^t \end{aligned} \quad (43)$$

where the second line in (43) follows from Euler's Theorem on homogeneous functions which (using our assumption that $f(\boldsymbol{q})$ is linearly homogeneous in $\boldsymbol{q}$) implies that $f(\boldsymbol{q}^t) = \boldsymbol{q}^t \cdot \nabla f(\boldsymbol{q}^t) = \sum_{n=1}^{N} q_{tn} \partial f(\boldsymbol{q}^t)/\partial q_n$ for $t = 1, \ldots, T$. Substitute λ_t defined by (43) into equations (41) and after a bit of rearrangement, we obtain the following system of estimating equations:

$$\boldsymbol{p}^t = e^t \nabla f(\boldsymbol{q}^t)/f(\boldsymbol{q}^t); \qquad t = 1, \ldots, T. \quad (44)$$

The above equations assume that all products were purchased in each period t. However, equations (44) can be generalized to deal with the case of missing products. When product n

[*26] See Hicks (1940; 140)[47] on the concept of a reservation price.

[*27] It is workable if the functional form for the unit cost function is a CES (Constant Elasticity of Substitution) function because the reservation prices are known (and equal plus infinity); see Feenstra (1994)[35].

is missing in period t, we simply set q_{tn} equal to 0 and drop product n from the utility maximization problem defined by (40). This leads to the smaller system of estimating equations defined by (45):[*28]

$$p_{tn} = e^t[\partial f(\boldsymbol{q}^t)/\partial q_n]/f(\boldsymbol{q}^t); \qquad t = 1, \ldots, T; n \in S(t). \tag{45}$$

Equations (45) define a *system of inverse demand functions*. We could assume a suitable functional form for the utility function $f(\boldsymbol{q})$, add error terms of the right hand sides of these equations and use the resulting system of equations as estimating equations to determine the unknown parameters that characterize the function $f(\boldsymbol{q})$.[*29] We also require a normalization on the parameters that define $f(\boldsymbol{q})$ in order to obtain a unique function.

It is usual in estimating systems of consumer demand equations to assume no missing prices and also to assume that the errors in the N equations pertaining to a single period are correlated so that a variance covariance matrix with $N(N+1)/2$ unknown parameters is also estimated. In our present context where we have 80 products, this strategy becomes unworkable. One strategy to solve this problem is to stack the estimating equations into a single estimating equation with only one variance parameter to deal with. However, this problem runs into a difficulty for National Statistical Offices: in general, the resulting parameter estimates are not invariant to the units in which we measure the products.[*30] Thus the resulting price and quantity indexes will also not be invariant to changes in the units of measurement. A solution to these problems is to switch from prices as the dependent variables to expenditure shares. Thus multiply both sides of equation tn in equations (45) by q_{tn} and divide by period t expenditure e^t. This leads to the nth expenditure share in period t, s_{tn}, as the dependent variable. These operations lead to the following system of *inverse demand share estimating equations* where e_{tn} is an error term:[*31]

$$s_{tn} = q_{tn}[\partial f(\boldsymbol{q}^t)/\partial q_n]/f(\boldsymbol{q}^t) + e_{tn}; \qquad t = 1, \ldots, T; n \in S(t). \tag{46}$$

When product n in period t is not available, $s_{tn} = q_{tn} = 0$ so equations (46) are valid for $t = 1, \ldots, T$ and $n = 1, \ldots, N$. However, note that the error term e_{tn} is equal to 0 when $q_{tn} = 0$. We stacked the resulting augmented equations (46) into a single estimating equation. In particular, rather than specifying an explicit error structure for equations (46), we assumed that the unknown parameters which characterize the chosen utility function $f(\boldsymbol{q})$ are estimated by solving the nonlinear least squares minimization problem (47) below with respect to the choice of these parameters:[*32]

$$\min_{\text{parameters of } f(q)} \textstyle\sum_{t=1}^{T} \sum_{n=1}^{N} \{s_{tn} - [q_{tn} f_n(\boldsymbol{q}^t))/f(\boldsymbol{q}^t)]\}^2 \tag{47}$$

[*28] For products n which are missing in period t, we can use the equation $p^*_{tn} \equiv e^t[\partial f(\boldsymbol{q}^t)/\partial q_n]/f(\boldsymbol{q}^t)$, $n \notin S(t)$, to define the Hicksian reservation price p^*_{tn} for that product.

[*29] Equations (45) are equivalent to the equations $f_n(\boldsymbol{q}^t)/p_{tn} = f(\boldsymbol{q}^t)/e^t \equiv \mu_t$ for $n \in S(t)$ and $t = 1, \ldots, 144$. When we add error terms to these equations, these equations tell us that in each period t, the marginal utility of a product divided by its price should be approximately equal to the same number μ_t, which turns out to equal $f(\boldsymbol{q}^t)/e^t$ which in turn is equal to $1/P^t$, where P^t is the aggregate price level for period t. Thus equations (45) are equations which fall out of the period t utility maximization problem. If the error terms are nonzero and large in magnitude, then this tells us that the assumption of approximate utility maximizing behaviour on the part of purchasers is probably not a good one.

[*30] For our particular application to rice products, this was not a problem since the package size was held constant across rice products.

[*31] Note that the right hand side of equation tn in equations (46) is an elasticity and so it should be invariant to changes in the units of measurement. The left hand side of equation tn in (46) is an expenditure share which is also invariant to changes in the units of measurement.

[*32] Our estimating method is not perfect from an econometric perspective since we have included observations in the minimization problem defined by (50) where $q_{tn} = s_{tn} = 0$. We have also neglected the fact that the shares s_{tn} sum to 0 in each region-period t; i.e., our estimation procedure does not take

where $f_n(\boldsymbol{q}^t) \equiv \partial f(\boldsymbol{q}^t)/\partial q_n$. A normalization on the parameters which characterize $f(\boldsymbol{q})$ is also required in order to obtain unique parameter estimates.

Once the unknown parameters characterizing $f(\boldsymbol{q})$ have been estimated, we can calculate period t aggregate quantities Q^t and the corresponding price levels P^t using the following definitions:

$$Q^t \equiv f(\boldsymbol{q}^t); P^t \equiv e^t/f(\boldsymbol{q}^t); \qquad t = 1, \ldots, T. \tag{48}$$

Note that the resulting quantity levels Q^t will satisfy the strong identity test for quantities: if $\boldsymbol{q}^r = \boldsymbol{q}^t$, then $f(\boldsymbol{q}^r) = f(\boldsymbol{q}^t)$, and hence $Q^r = Q^t$.[*33]

The bottom line is this: it is virtually impossible to estimate systems of direct consumer demand functions when there are missing prices but it is reasonably straightforward to estimate systems of inverse demand functions. It is possible to estimate the utility function directly when there are missing prices but very difficult to estimate the corresponding dual unit cost function.

In the following three sections, we will work through the algebra presented in this section for three specific functional forms for $f(\boldsymbol{q})$.

7 The Econometric Estimation of Linear Preferences

Our first example of the methodology explained in the previous section is the case where the utility function is a homogeneous linear function of the quantities consumed. Thus we assume that $f(\boldsymbol{q}, \boldsymbol{\alpha})$ has the following functional form:

$$f(\boldsymbol{q}, \boldsymbol{\alpha}) \equiv \textstyle\sum_{n=1}^{N} \alpha_n q_n = \boldsymbol{\alpha} \cdot \boldsymbol{q}. \tag{49}$$

The least squares minimization problem (49) becomes the following problem:

$$\min_{\alpha' s} \textstyle\sum_{t=1}^{T} \sum_{n=1}^{N} \{s_{tn} - [q_{tn}\alpha_n/\boldsymbol{\alpha} \cdot \boldsymbol{q}^t]\}^2. \tag{50}$$

If $\boldsymbol{\alpha}^* \equiv [\alpha_1^*, \ldots, \alpha_N^*]$ is a solution to (50), then it can be seen that $\lambda\boldsymbol{\alpha}^*$ is also a solution to (50) where λ is any positive number. This non-uniqueness always occur when we attempt to estimate utility functions. The scale of utility is arbitrary so we need to impose at least one normalization on the estimated parameters in order to obtain a cardinal measure of utility.

There is another possible problem with the minimization problem defined by (50): it can be the case that there is no solution to (50). For example, suppose that there are only 2 periods and 2 products in scope. Suppose further that product 1 is only available in period 1 and product 2 is only available in period 2. In this case, there are only 2 independent estimating equations for the nonlinear minimization problem defined by (50):

$$1 = \alpha_1 q_{11}/(\alpha_1 q_{11} + \alpha_2 0) = 1; \tag{51}$$

$$1 = \alpha_2 q_{22}/(\alpha_1 0 + \alpha_2 q_{22}) = 1. \tag{52}$$

into account the fact that $\sum_{n=1}^{80} s_{tn} = 0$ for $t = 1, \ldots, 144$. However, it is convenient to work with the class minimization problems defined by (50) for specific functional forms for $f(\boldsymbol{q})$ because it enables us to deduce axiomatic properties of the resulting price and quantity indexes; i.e., see Diewert (2005)[21] (2023)[25].

*33 The corresponding period t price level P^t regarded as a function of $\boldsymbol{p}^t$ will be linearly homogeneous in the components of $\boldsymbol{p}^t$.

It can be seen that it is not possible to obtain estimates for the quality adjustment parameters α_1 and α_2 in this situation. We need some product overlap between the periods in order to obtain solutions to (50).

In order to solve the problems of non-uniqueness and non-existence in general, we assume that there is a product that is present in all 144 periods and we assume that each product in scope is purchased in at least one period. In our rice products data set, there were 11 products that were present in all 144 region-months. Product 4 was the lowest number product that was present in all periods so we set $\alpha_4 = 1$.

We used the nonlinear regression option in Shazam (see White (2004)[69]) to solve (50) with $\alpha_4 = 1$. The program ran for 87 iterations and took 80.6 seconds to converge. The R^2 between observed and predicted prices was 0.9964. The final loglikelihood was 56774.19. All of the estimated α_n turned out to be positive (and reasonable). Note that the R^2 for the Weighted Time Product Dummy Model (another linear preferences model) was 0.9941. Denote the estimated α_n by α_n^* except define $\alpha_4^* \equiv 1$. Define the vector $\boldsymbol{\alpha}^* \equiv [\alpha_1^*, \ldots, \alpha_{80}^*]$ and define preliminary quantity and price levels, Q^{t^*} and P^{t^*} for period t (a region-month), as follows:

$$Q^{t^*} \equiv \boldsymbol{\alpha}^* \cdot \boldsymbol{q}^t; P^{t^*} \equiv e^t/Q^{t^*}; \qquad t = 1, \ldots, 144. \tag{53}$$

Normalize the sequence of price levels P^{t^*} into the series $P^{t^{**}}$ which is such that the normalized sequence of price levels equals 1 for $t = 25$ (month 1 for Tokyo):

$$P^{t^{**}} \equiv P^{t^*}/P^{25^*}; \qquad t = 1, \ldots, 144. \tag{54}$$

Finally define the econometric *linear utility price levels* for regions 1-6 for $m = 1, \ldots, 24$ as follows:

$$P_{\text{LU}}^{1,m} \equiv P^{m^{**}}; P_{\text{LU}}^{2,m} \equiv P^{(24+m)^{**}}; P_{\text{LU}}^{3,m} \equiv P^{(48+m)^{**}}; P_{\text{LU}}^{4,m} \equiv P^{(72+m)^{**}}; P_{\text{LU}}^{5,m} \equiv P^{(96+m)^{**}};$$
$$P_{\text{LU}}^{6,m} \equiv P^{(120+m)^{**}}. \tag{55}$$

The regional Linear Utility Price indexes, P_{LU}^1-P_{LU}^6 are listed in Table 4 of the Appendix. These indexes were stacked into the index P_{LU}^t which is plotted on Chart 3 below.

The Linear Preferences price indexes, P_{LU}^t, are generally higher than the other indexes and very much higher for the smaller population Prefectures, which are Tottori (observations 73-96), Kochi (observations 97-120) and Kagoshima (observations 121-144). The Geary Khamis indexes, P_{GK}^t, tended to be lower than the other indexes, particularly for the smaller population Prefectures. All four indexes were very close to each other for the highest population Prefecture, Tokyo (observations 25-48). The Implicit Weighted Time Product Dummy indexes, P_{IWTPD}^t, and the GEKS indexes, P_{GEKS}^t, were generally in the middle and fairly close to each other. It is interesting that the first 3 indexes are all consistent with linear preferences but they turned out to be quite different for our particular data set. What is striking is the fact that price levels in the 3 lowest population Prefectures (Tottori, Kochi and Kagoshima) were generally much higher than price levels in the first 3 higher population Prefectures (Hokkaido, Tokyo and Kyoto). These differences indicate that there may be a problem in using national price indexes in order to deflate consumer expenditures into real consumption aggregates since national Consumer Price Indexes do not take differing interregional price levels into account in their construction. Thus poverty measures and measures of national real consumption may be inaccurate to a significant degree.

Could these Linear Utility Price indexes that are based on the estimation of a linear utility function be acceptable to National Statistical Offices? It seems that they might be acceptable

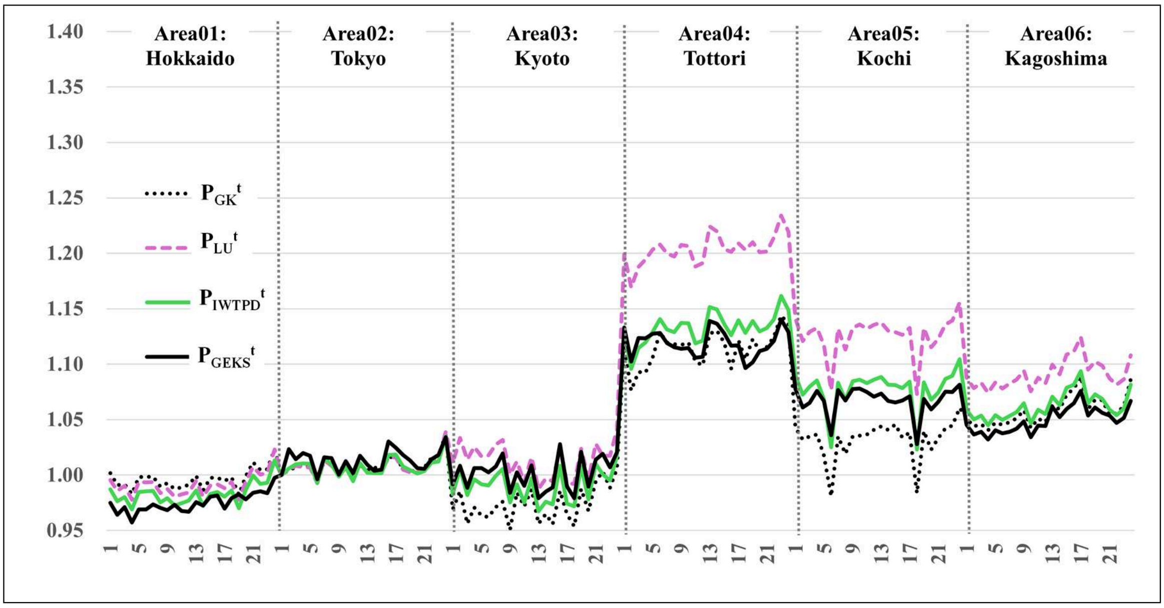

Chart 3 Geary Khamis, Linear Preferences,Implicit Weighted TPD and GEKS Price Indexes

since NSOs are already estimating linear preferences when they use Geary (1958)[41] Khamis (1970)[53] multilateral indexes to construct portions of their CPIs. Many offices also use hedonic time dummy regression models to quality adjust products that quickly appear and then disappear. As we have seen in section 5 above, hedonic regression models that use time dummy variables are also based on the (implicit) assumption of linear preferences.

8 The Estimation of CES Preferences

Our second example of the methodology explained in section 6 is the case where the utility function is a CES (Constant Elasticity of Substitution) function, $f(\boldsymbol{q})$ defined as follows in the case of N products:[*34]

$$f(\boldsymbol{q}) \equiv [\sum_{n=1}^{N} \alpha_n (q_n)^k]^{1/k} \quad (56)$$

where the α_n are positive parameters and the parameter k satisfies the following inequalities:

$$0 < k \leq 1.^{*35} \quad (57)$$

Note that if the parameter k equals 1, then the CES utility function defined by (56) becomes the linear utility function that was discussed in the previous section.

Recall from section 6 that the observed period t vector $\boldsymbol{q}^t$ solves the period t utility maximization problem if $p_{tn} = e^t[\partial f(\boldsymbol{q}^t)/\partial q_n]/f(\boldsymbol{q}^t)$ for all $n \in S(t)$. If we multiply both sides of

[*34] In the mathematics literature, if the α_n sum to one, this aggregator function or utility function is known as a power mean or a mean of order k; see Hardy, Littlewood and Pólya (1934; 12-13)[44]. This functional form was popularized by Arrow, Chenery, Minhas and Solow (1961)[1] in the context of production theory. For more on estimating CES utility functions, see Balk (1999)[6], Melser (2006)[56], Diewert (2020a)[23] (2020b)[24] and de Haan and Krsinich (2024)[16].

[*35] We require that $k \leq 1$ to ensure that the utility function is concave in the components of $\boldsymbol{q}$ and we require that $k > 0$ in order to ensure that the utility function is well defined if any component of the $\boldsymbol{q}^t$ vector happens to be equal to 0. The restrictions $0 < k < 1$ are also required in order to apply Feenstra's (1994)[35] methodology for measuring the welfare effects of increased (or decreased) product choice.

equation n by q_{tn}, then these first order necessary conditions become the following estimating equations:

$$s_{tn} = p_{tn}q_{tn}/e^t = q_{tn}[\partial f(\boldsymbol{q}^t)/\partial q_n]/f(\boldsymbol{q}^t) = \alpha_n(q_{tn})^k/\textstyle\sum_{i\in S(t)}\alpha_i(q_{ti})^k;$$
$$t = 1,\ldots,T; n \in S(t). \tag{58}$$

If $q_{tn} = 0$, then $s_{tn} = 0$. Thus the equations (58) can be replaced with the following equations:

$$s_{tn} = \alpha_n(q_{tn})^k/\textstyle\sum_{i=1}^N\alpha_i(q_{ti})^k; \qquad t = 1,\ldots,T; n = 1,\ldots,N. \tag{59}$$

In our particular case, $N = 80$ and $T = 144$. We obtained estimates for the CES utility function by solving the following nonlinear least squares minimization problem:

$$\min_{\alpha' s}\textstyle\sum_{t=1}^{144}\sum_{n=1}^{80}\{s_{tn} - [\alpha_n(q_{tn})^k/\sum_{i=1}^N\alpha_i(q_{ti})^k]\}^2. \tag{60}$$

Note that if $q_{tn} = 0$, then both s_{tn} and $\alpha_n(q_{tn})^k$ equal zero. If $\boldsymbol{\alpha}^* \equiv [\alpha_1^*,\ldots,\alpha_N^*]$ and k is a solution to (60), then it can be seen that $\lambda\boldsymbol{\alpha}^*$ and k is also a solution to (60) where λ is any positive number. Thus we imposed the normalization $\alpha_4 = 1$ because product 4 was the first product on our list of products that was present in all 144 region-periods.

We used the nonlinear regression option in Shazam (see White (2004)[69]) to solve (60) with α_4 set equal to 1. The starting values for the unknown α_n were the final estimated coefficients from the linear model estimated in the previous section. Our starting value for the parameter k was 1. The starting log likelihood for the present nonlinear regression defined by (60) was almost equal to the final log likelihood for the linear model,[*36] which is a good check on our estimating code. The program ran for 91 iterations and took 8.6 minutes to converge. The R^2 between observed and predicted prices was 0.9967. The final log likelihood was 57333.19, a gain of 559.00 over the linear preferences log likelihood for adding 1 parameter. All of the estimated α_n turned out to be positive. Recall that the R^2 for the linear preferences model was 0.9964. The estimated k was $k^* = 0.95668$ with an estimated standard error equal to 0.00125. The corresponding elasticity of substitution σ^* was equal to:[*37]

$$\sigma^* \equiv 1/(1 - k^*) = 23.086. \tag{61}$$

Denote the estimated α_n by α_n^* and define $\alpha_4^* \equiv 1$. Define the vector $\boldsymbol{\alpha}^* \equiv [\alpha_1^*,\ldots,\alpha_{80}^*]$ and define preliminary CES quantity and price levels, Q^{t^*} and P^{t^*} for period t (a region-month), as follows:

$$Q^{t^*} \equiv [\textstyle\sum_{n=1}^N\alpha_n^*(q_{tn})^{k^*}]^{1/k^*}; P^{t^*} \equiv e^t/Q^{t^*}; \qquad t = 1,\ldots,144. \tag{62}$$

[*36] The final log likelihood for the linear model was 56774.19; the starting likelihood for the CES model was 56773.57.

[*37] For the definition of the elasticity of substitution and derivation of the formula (61), see Diewert(2020a; 31-37)[23]. Our estimated elasticity of substitution is high compared to many estimates in the literature. The reason for this is that we estimated the CES utility function directly (and the companion unit costs indirectly) whereas most of the literature estimates the dual CES unit cost function. The CES unit cost function is only one parameter away from the linear cost function which is dual to a Leontief (no substitution) utility function. The CES utility function is only one parameter away from a linear utility function. At the first level of aggregation, a linear utility function is more likely to provide a much better approximation to preferences than a linear unit cost function. For our particular data set, the CES direct utility function model fit the data much better than the corresponding CES unit cost function as we shall see. If there were no errors in fit in either model, we would get the same result but of course, neither model fits the data perfectly.

Note that the P^{t^*} are defined *indirectly* using the product test, $P^{t^*}Q^{t^*} = e^t$. Normalize the sequence of price levels P^{t^*} into the series P^t_{CES} which is such that the normalized sequence of price levels equals 1 for $t = 25$ (month 1 for Tokyo):

$$P^t_{\text{CES}} \equiv P^{t^*}/P^{25^*}; \qquad t = 1, \ldots, 144. \tag{63}$$

Finally define the econometric *CES utility function price levels for regions 1-6* as follows:

$$\begin{aligned} &P^{1,m}_{\text{CES}} \equiv P^m_{\text{CES}}; P^{2,m}_{\text{CES}} \equiv P^{(24+m)}_{\text{CES}}; P^{3,m}_{\text{CES}} \equiv P^{(48+m)}_{\text{CES}}; \\ &P^{4,m}_{\text{CES}} \equiv P^{(72+m)}_{\text{CES}}; P^{5,m}_{\text{CES}} \equiv P^{(96+m)}_{\text{CES}}; P^{6,m}_{\text{CES}} \equiv P^{(120+m)}_{\text{CES}}; \qquad m = 1, \ldots, 24. \end{aligned} \tag{64}$$

The CES indexes, P^t_{CES}, are plotted on Chart 4 below and the regional CES indexes $P^{1,m}_{\text{CES}}$ - $P^{6,m}_{\text{CES}}$ are listed in Table 5 in the Appendix.

Our CES price indexes P^t_{CES} were defined *indirectly* using the estimated utility levels to deflate actual expenditure levels into aggregate price levels. There is another indirect method that could be used to define CES price levels given that we have estimated the CES utility function: we could use the estimated utility function to solve the following period t unit cost minimization problem for each period t:

$$\min_q\{\textstyle\sum_{n\in S(t)} p_{tn}q_n : f(q_1, q_2, \ldots, q_{80}) \geq 1; q_n = 0 \text{ if } n \notin S(t)\} = c(\boldsymbol{p}^t); \qquad t = 1, \ldots, 144. \tag{65}$$

Suppose for the moment that there are no missing products for the period t cost minimization problem defined by (65) and our $f(\boldsymbol{q})$ is defined by (62); i.e., $f(\boldsymbol{q}) \equiv [\sum_{n=1}^N \alpha^*_n (q_n)^{k^*}]^{1/k^*}$. Then it can be shown that the CES unit cost function has the following functional form:[*38]

$$c(\boldsymbol{p}^t) = [\textstyle\sum_{n=1}^N \beta^*_n (p_{tn})^{\kappa^*}]^{1/\kappa^*} \tag{66}$$

where the parameters β^*_n and κ^* are defined as follows:

$$\beta^*_n \equiv (\alpha^*_n)^{1/(1-k^*)} \text{ for } n = 1, ..., 80 \text{ and } \kappa^* \equiv -k^*/(1-k^*) = -22.0856.^{*39} \tag{67}$$

In order to deal with the case where some products are not available in period t, Feenstra (1994)[35] assumed that the parameter κ^* which appears in definition (66) satisfies $\kappa^* < 0$. This allowed Feenstra to set the reservation prices for the missing products equal to $+\infty$ and thus when $\kappa^* < 0$, an infinite price p_{tn} raised to a negative power generates a zero; i.e., if product n is unavailable in period t, then $(p_{tn})^{\kappa^*} = 1/(+\infty)^{|\kappa^*|} = 0$. Thus with infinite reservation prices for missing products, period t unit cost is equal to:

$$c(\boldsymbol{p}^t) = [\textstyle\sum_{n=1}^N \beta^*_n (p_{tn})^{\kappa^*}]^{1/\kappa^*} = [\sum_{n\in S(t)} \beta^*_n (p_{tn})^{\kappa^*}]^{1/\kappa^*} \equiv P^{t^*}; \qquad t = 1, \ldots, 144 \tag{68}$$

where the β^*_n and κ^* are defined by (67). Normalize the resulting period t unit costs P^{t^*} into the following Alternative CES price levels, $P^t_{\text{ACES}} = P^{t^*}/P^{25^*}$ for $t = 1, \ldots, 144$. These alternative indirectly derived CES price indexes are plotted on Chart 4 below. Counterparts to definitions (64) are used to decompose P^t_{ACES} into the 6 regional indexes, $P^{1,m}_{\text{ACES}}$ – $P^{6,m}_{\text{ACES}}$, which are listed in Table 5 in the Appendix. It should be noted that the use of the definitions in (67) led to β^*_n that were tiny if $\alpha^*_n > 1$ or β^*_n that were huge if $\alpha^*_n < 1$. This in turn led

[*38] See for example Diewert (2020a; 37)[23].
[*39] Recall that our estimated k was $k^* = 0.95668$.

to estimated unit costs which exhibited excessive fluctuations; see Chart 4. This method for forming CES price indexes is not recommended if the parameter κ^* is large in magnitude or if the elasticity of substitution $\sigma \equiv 1 - \kappa^*$ is large.

Our final CES set of regional price indexes is obtained by directly estimating the unit cost function defined by (68).[*40] Shephard's Lemma can be used to obtain cost minimizing quantities as functions of prices when preferences are represented by a differentiable unit cost function.[*41] Thus if preferences are represented by the CES utility function that is dual to a CES unit cost function that is defined by (66) in the case of no missing products, then $q_{tn} = Q^t \partial c(\boldsymbol{p}^t)/\partial p_n$ for $n = 1, \ldots, N$. This approach can be generalized to the case of missing products and it leads to the following system of estimating equations:

$$s_{tn} = \beta_n (p_{tn})^\kappa / \textstyle\sum_{j \in S(t)} \beta_j (p_{tj})^\kappa + e_{tn}; \qquad t = 1, \ldots, 144; n \in S(t). \tag{69}$$

It proved technically difficult to set up the nonlinear least squares minimization problem that is associated with equations (69) so we used the following approach that is often used in the literature: take logarithms of both sides of equations (69) for $n \in S(t)$, subtract the resulting logarithmic equation for product 4 in period t from the corresponding $\log s_{tn}$ equation and set $\beta_4 = 1$ (so that the logarithm of β_4 equals 0). We obtain the following system of estimating equations where $\alpha_n \equiv \ln(\beta_n)$ for $n = 1, \ldots, 80$ and e_{tn} is an error term:

$$y_{tn} = \alpha_n + \kappa x_{tn} + e_{tn}; \qquad t = 1, \ldots, 144; n \in S(t) \tag{70}$$

where the e_{tn} are error terms, $y_{tn} \equiv \ln s_{tn} - \ln s_{t4}$ and $x_{tn} \equiv \ln p_{tn} - \ln p_{t4}$ for $t = 1, \ldots, 144; n \in S(t)$.

The parameters κ and α_n for $n = 1, \ldots, 80$ which appear in equations (70) were estimated in a single stacked linear regression. The resulting R^2 between observed and predicted was only 0.7289. Recall that the R^2 for the direct estimation of the CES utility function (using a stacked regression) was 0.9967 and the R^2 for the direct estimation of a linear utility function was 0.9964. The sum of absolute errors for the present regression was 5285.3 whereas the sum of absolute errors for the direct estimation of the CES utility function was 7.2587. This is a very large difference in fit. Our new estimate for the parameter κ was $\kappa^* = -5.4994$ with a standard error equal to 0.1036. Thus our new estimate for the elasticity of substitution is:

$$\sigma^* \equiv 1 - \kappa^* = 6.4994. \tag{71}$$

Recall that our earlier estimate for the elasticity of substitution was equal to 23.086. This is a huge difference.

Denote the estimated α_n by α_n^* and define $\alpha_4^* \equiv 1$. Define the vector $\boldsymbol{\alpha}^* \equiv [\alpha_1^*, \ldots, \alpha_{80}^*]$ and define preliminary *cost function based* CES quantity and price levels, Q^{t^*} and P^{t^*} for period t (a region-month), as follows:

$$P^{t^*} \equiv [\textstyle\sum_{n \in S(t)} \alpha_n^* (p_{tn})^{\kappa^*}]^{1/\kappa^*}; Q^{t^*} \equiv e^t / P^{t^*}; \qquad t = 1, \ldots, 144. \tag{72}$$

Note that the Q^{t^*} are defined *indirectly* using the product test, $P^{t^*} Q^{t^*} = e^t$. Normalize the sequence of cost function based price levels P^{t^*} into the series P^t_{CCES} which is such that the

[*40] There are many other methods that have appeared in the literature for estimating CES price indexes; see for example Feenstra (1994)[35], Balk (1999)[6], Diewert and Feenstra (2017)[27] (2022)[28], Diewert (2020a)[23] and de Haan and Krsinich (2024)[16].

[*41] See Samuelson and Swamy (1974)[62] or Diewert (1974)[17] (1976)[18] for the details on how this dual approach works.

normalized sequence of price levels equals 1 for $t = 25$ (month 1 for Tokyo):

$$P^t_{\mathrm{CCES}} \equiv P^{t^*}/P^{25^*}; \qquad t = 1, \ldots, 144. \tag{73}$$

Finally define the econometric *Cost Function Based CES utility function price levels for regions 1-6* as $P^{r,m}_{\mathrm{CCES}}$ using P^t_{CCES} defined by (73) and an appropriate modification of definitions (64). The Cost Based CES indexes, P^t_{CCES}, are plotted on Chart 4 below and the regional CCES indexes $P^{1,m}_{\mathrm{CCES}}$ - $P^{6,m}_{\mathrm{CCES}}$ are listed in Table 5 in the Appendix.

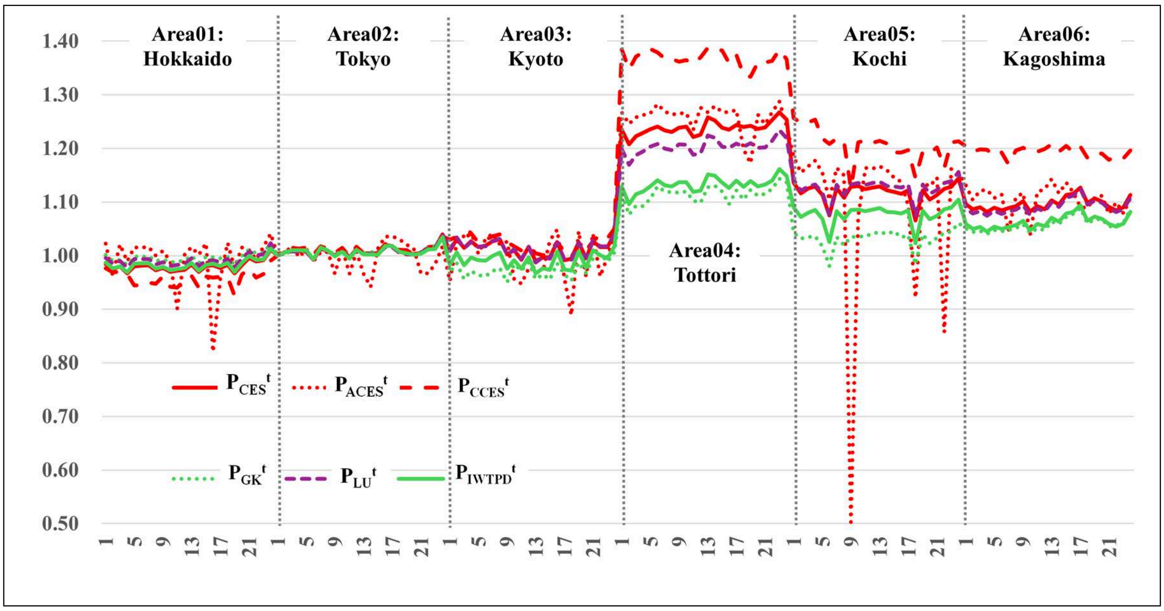

Chart 4 Alternative CES, GK, Linear Utility and Implicit Weighted TPD Price Indexes

The six alternative price indexes for Tottori (observations 73-96) are widely separated with almost 30 percentage points difference between the highest and lowest index. The ordering of the indexes from high to low for the Tottori observations is as follows. The cost function based CES price index P^t_{CCES} is highest, followed by the unit cost function price levels P^t_{ACES} that were obtained by using the estimated CES utility function parameters to solve for the dual unit costs, followed by our first CES estimates for price levels P^t_{CES} that were obtained by deflating expenditures by CES utility levels, followed by the linear utility function price levels P^t_{LU}. The two lowest series were the Geary Khamis price indexes P^t_{GK} and the Weighted Implicit Time Product Dummy indexes P^t_{WITPD}. For Prefectures 4-6, the GK Price indexes tended to be lowest. It is clear that P^t_{ACES} is not a suitable index due to its extreme volatility. A linear utility function is likely to overstate substitution possibilities so it is not surprising that the more flexible CES based price indexes, P^t_{CES} and P^t_{CCES}, are generally higher than the linear utility function price indexes P^t_{LU}. What is surprising is that the cost function based CES indexes P^t_{CCES} are so much higher than the utility function based CES price indexes P^t_{CES}. Since the utility function based indexes P^t_{CES} fit the data so much better than the cost function based indexes P^t_{CCES}, the former indexes are preferred.

With estimates for the elasticity of substitution in hand, we can use Feenstra's 1994 methodology to estimate the effects on welfare of different degrees of product availability across the 6 Prefectures. Suppose that the CES unit cost function for month m in region r is defined as $c(p^{r,m}) \equiv [\sum_{n \in S(r,m)} \alpha_n (p_{rmn})^{\kappa}]^{1/\kappa}$ where $S(r,m)$ is the set of rice products n that are purchased in month m in region r and the parameter κ is less than 0. The unit cost $c(p^{r,m})$ represents the rice price level for region r in month m. Thus the rice consumer price index for month m in region r relative to the price level in Tokyo for the same month m is the ratio of unit costs, $c(p^{r,m})/c(p^{2,m})$. Feenstra obtained the following decomposition of $c(p^{r,m})/c(p^{2,m})$:

$$\begin{aligned} P_{\text{CES}}^{r,m}/P_{\text{CES}}^{2,m} &\equiv c(p^{r,m})/c(p^{2,m}); \qquad r=1,\ldots,6; m=1,\ldots,24 \\ &\equiv [\textstyle\sum_{n\in S(r,m)}\alpha_n(p_{rmn})^{\kappa}]^{1/\kappa}/[\sum_{n\in S(2,m)}\alpha_n(p_{2mn})^{\kappa}]^{1/\kappa} \\ &= [A^{r,m}]\times[B^{r,m}]\times[C^{r,m}] \end{aligned} \tag{74}$$

where the three indexes in the last line of equations (74) are defined as follows:[*42]

$$A^{r,m} \equiv [\textstyle\sum_{n\in S(r,m)\cap S(2,m)}\alpha_n(p_{rmn})^{\kappa}]^{1/\kappa}/[\sum_{n\in S(r,m)\cap S(2,m)}\alpha_n(p_{2mn})^{\kappa}]^{1/\kappa}; \quad r=1,\ldots,6; m=1,\ldots,24 \tag{75}$$

$$B^{r,m} \equiv [\textstyle\sum_{n\in S(r,m)}\alpha_n(p_{rmn})^{\kappa}]^{1/\kappa}/[\sum_{n\in S(r,m)\cap S(2,m)}\alpha_n(p_{rmn})^{\kappa}]^{1/\kappa}; \quad r=1,\ldots,6; m=1,\ldots,24 \tag{76}$$

$$C^{r,m} \equiv [\textstyle\sum_{n\in S(r,m)\cap S(2,m)}\alpha_n(p_{2mn})^{\kappa}]^{1/\kappa}/[\sum_{n\in S(2,m)}\alpha_n(p_{2mn})^{\kappa}]^{1/\kappa}; \quad r=1,\ldots,6; m=1,\ldots,24. \tag{77}$$

The left hand side of (74) is $P_{\text{CES}}^{r,m}/P_{\text{CES}}^{2,m} \equiv c(p^{r,m})/c(p^{2,m})$ which is the overall CES rice Cost of Living index for Region r relative to Toyko in month m. The index $A^{r,m}$ is the relative cost of achieving the same utility level if purchasers faced the prices of rice products that are common to both Regions r and Tokyo (region 2) in month m with the Region r price level in the numerator and the Region 2 prices in the denominator. The index $B^{r,m}$ has the Region r cost of attaining one unit of utility if purchasers faced the actual prices of month m in Region r in the numerator and the denominator is the hypothetical Region r cost of attaining one unit of utility if only products found in Regions r *and* 2 were available. Thus $B^{r,m} \le 1$. The lower $B^{r,m}$ is, the bigger is the benefit to purchasers in Region r of having extra products that were not available in Tokyo in month m. The Region r prices in month m are used in both numerator and denominator. The *reciprocal* of the index $C^{r,m}$ is again equal to the ratio of two unit costs: the cost of achieving one unit of utility in Region 2 in month m if region 2 products were available and the cost of achieving one unit of utility in Region 2 if purchasers faced Region 2 prices in month m but were restricted to purchasing products that were available in both Regions 2 and r in month m. Region 2 prices in month m are used in the numerator and denominator. Thus $1/C^{r,m} \le 1$ or $C^{r,m} \ge 1$ and the bigger $C^{r,m}$ is, the bigger is the benefit to Region 2 to having its choice set $S(2,m)$ relative to the more restricted choice set $S(2,m)\cap S(r,m)$. We define the month m net cost to Region r of having its choice set relative to the corresponding month m, Region 2 choice set to be the product of $B^{r,m}$ and $C^{r,m}$:[*43]

$$D^{r,m} \equiv [B^{r,m}]\times[C^{r,m}]; \qquad r=1,\ldots,6; m=1,\ldots,24. \tag{78}$$

If $D^{r,m} > 1$, then the difference in choice sets between Region r and Region 2 adds to the Region r cost of living.

[*42] When $r=2$, the indexes $A^{r,m}, B^{r,m}$ and $C^{r,m}$ are all equal to 1 for $m=1,\ldots,24$.

[*43] Of course, the product of $D^{r,m}$ and the matched product CES index $A^{r,m}$ is equal to the actual Cost of Living index between Regions r and 2 for month m, $P_{\text{CES}}^{r,m}/P_{\text{CES}}^{2,m}$. Thus $D^{r,m}$ can be interpreted as an adjustment to the matched product index that takes into account differences in product availability. For more details on the Feenstra methodology, see Feenstra (1994)[35], Balk (1999)[6], Melser (2006)[56], Diewert and Feenstra (2017)[27] (2022)[28] and Diewert (2020b, 41-44)[24].

Feenstra (1994)[35] showed that if Regions r and 2 have a product in common for month m, then it is possible to estimate the indexes $B^{r,m}$ and $C^{r,m}$ without estimating the CES cost function, provided that we have an estimate for the parameter κ or equivalently for the elasticity of substitution, $\sigma \equiv 1-\kappa$. His method starts by defining the following *observable expenditure or sales ratios*:

$$\lambda^{r,m} \equiv \textstyle\sum_{n\in S(r,m)} p_{rmn}q_{rmn} / \sum_{n\in S(r,m)\cap S(2,m)} p_{rmn}q_{rmn}; \quad r=1,\ldots,6; m=1,\ldots,24 \quad (79)$$

$$\mu^{r,m} \equiv \textstyle\sum_{n\in S(r,m)\cap S(2,m)} p_{2mn}q_{2mn} / \sum_{n\in S(2,m)} p_{2mn}q_{2mn}; \quad r=1,\ldots,6; m=1,\ldots,24. \quad (80)$$

$\lambda^{r,m}$ is the ratio of rice expenditures in Prefecture r in month m relative to rice expenditures in the same month restricted to the set of products that are available in *both* Prefecture r and Prefecture 2 (Tokyo). Thus this ratio must satisfy the inequality $\lambda^{r,m} \geq 1$. $\mu^{r,m}$ is the reciprocal of the ratio of rice expenditures in Prefecture 2 in month m relative to rice expenditures in the same month restricted to the set of products that are available in *both* Prefecture 2 and Prefecture r. Thus $\mu^{r,m}$ must satisfy the inequality $\mu^{r,m} \leq 1$. Of course, when $r=2$, it is the case that $\lambda^{r,m} = \mu^{r,m} = 1$ for $m=1,\ldots,24$. The expenditure ratios defined by $\lambda^{r,m}$ and $\mu^{r,m}$ are listed in Table 6 in the Appendix. Finally, Feenstra (1994)[35] showed that:

$$B^{r,m} = [\lambda^{r,m}]^{1/\kappa} \text{ and } C^{r,m} = [\mu^{r,m}]^{1/\kappa}; \qquad r=1,\ldots,6; m=1,\ldots,24. \quad (81)$$

Thus if κ (or the elasticity of substitution $\sigma = 1-\kappa$) is known or has been estimated, then $B^{r,m}$ and $C^{r,m}$ can readily be calculated as simple ratios of sums of observable expenditures raised to the power $1/\kappa$. Thus the measures of price level change due to changes in product availability in Prefecture r relative to Prefecture 2 for month m, $D^{r,m}$ defined as $B^{r,m}$ times $C^{r,m}$, can be calculated. We have two estimates for the elasticity of substitution, $\sigma = 23.0856$ from our first CES model that estimated the CES utility function and $\sigma = 6.4994$ from our CES model that estimated the CES unit cost function. These alternative estimates for the elasticity of substitution lead to the two alternative estimates for κ equal to -22.0856 and -5.4994. Thus we can generate two sets of indexes $D^{r,m}$ using these two estimates for κ and the above definitions. These two sets of indexes are denoted by $D^{r,m}_{\text{CES}}$ and $D^{r,m}_{\text{CCES}}$ and are listed on Table 7 in the Appendix. These indexes are stacked and plotted as D^{t}_{CES} and D^{t}_{CCES} on Chart 5.

Of course, D^{t}_{CES} and D^{t}_{CCES} are equal to 1 for $t=25,\ldots,48$ since these indexes compare availability of products in Prefecture $r=2$ with the same Prefecture (Tokyo). The average increase in rice prices in Hokkaido and Kyoto due to differences in the availability of products in these two Prefectures relative to availability in Tokyo was only 0.41 percentage points on average using the estimates for σ from our estimation of the CES utility function but this average estimate increased to 1.66 or 1.67 percentage points using the lower estimate of σ that came out of our estimation of the CES unit cost function. The increase in the cost of living index due to limited availability of products was much greater for the smaller population Prefectures. Here are the *average increases in cost* due to limited product availability generated by the two CES models for the smaller Prefectures in percentage points: Tottori: 5.87 and 25.76; Kochi: 3.94 and 16.82; Kagoshima: 3.24 and 13.68. Thus for the smaller Prefectures, the cost function based indexes D^{t}_{CCES} lie far above the utility function based indexes D^{t}_{CES}. It can be seen that it is very important to obtain accurate and realistic estimates for the elasticity of substitution when applying the Feenstra methodology. Since our utility function based method for estimating the elasticity of substitution fit the data far better than the cost function based method, the smaller estimates for the increase in the cost of living due to smaller choice sets are our preferred estimates.

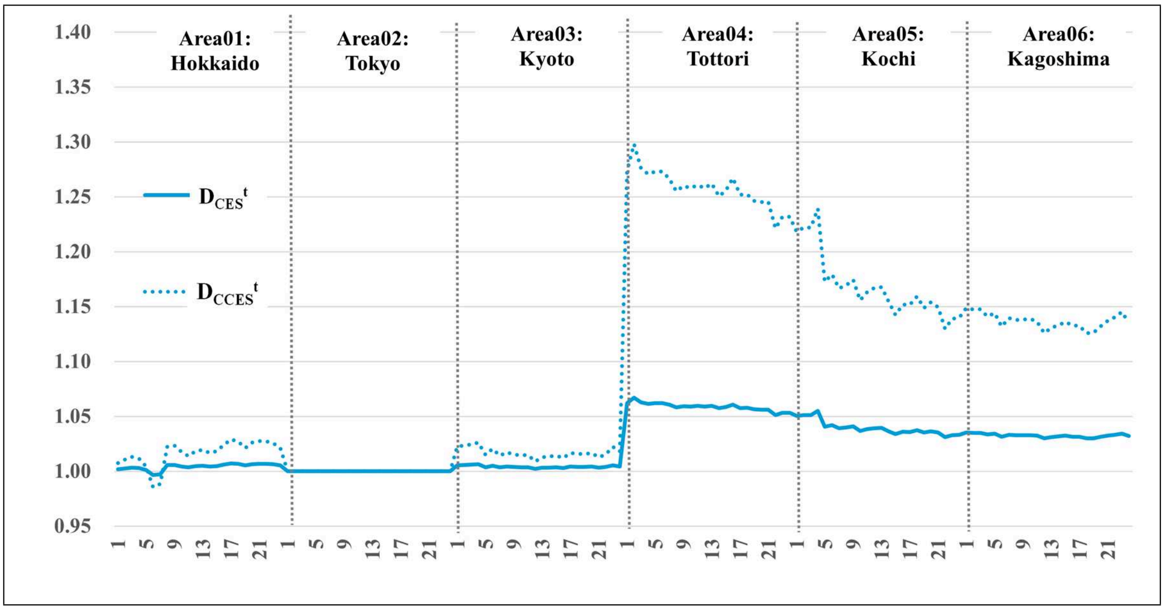

Chart 5 Two Measures of the Increase in the Price Level of Six Prefectures Relative to Tokyo Prefecture due to Differences in Product Availability

A problem with the CES functional form is that there is only one parameter, the elasticity of substitution σ, that is used to describe substitution possibilities between every pair of products. In the following section, we estimate a functional form for the purchaser's utility function that has a separate parameter to describe substitution possibilities for each product.

9 The Econometric Estimation of KBF Preferences with a Rank One Substitution Matrix

Konüs and Byushgens (1926)[54][*44] introduced the following functional form for a linearly homogeneous utility function:

$$f(\boldsymbol{q}) \equiv (\boldsymbol{q}^T \cdot \mathbf{A}\boldsymbol{q})^{1/2} = (\textstyle\sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}q_iq_j)^{1/2}; a_{ij} = a_{ji}; \qquad 1 \leq i \leq j \leq N. \qquad (82)$$

Thus $\mathbf{A}$ is an N by N symmetric matrix that contains $(N+1)N/2$ unknown a_{ij} parameters. The matrix $\mathbf{A}$ satisfies certain restrictions which are spelled out in Diewert (1976)[18]. Konüs and Byushgens and Diewert showed that this utility function is exact for the Fisher (1922)[36] Ideal quantity and price indexes so we call preferences defined by (82) KBF preferences.

Using the utility maximization framework which was described in section 6, the possible estimating equations (45) become the following *system of inverse demand functions*:[*45]

$$p_{tn} = e^t d_{tn}(\textstyle\sum_{j=1}^{N} a_{nj}q_{tj})/(\sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}q_{ti}q_{tj}); \quad t = 1, \ldots, 144; n = 1, \ldots, 80. \qquad (83)$$

where the dummy variable is defined as before; i.e., $d_{tn} \equiv 1$ if $n \in S(t)$ and define $d_{tn} \equiv 0$ if product n is not available in period t for $t = 1, \ldots, 144$ and $n = 1, \ldots, 80$.

[*44] See Diewert and Zelenyuk (2024)[33] for a translation and commentary on their paper.

[*45] We did not use these equations as our estimating equations because the resulting price and quantity indexes are not invariant to changes in the units of measurement.

We will not attempt to estimate all $(N+1)N/2$ unknown parameters a_{ij} in the KBF utility function defined by (82). In order to reduce the number of parameters in the $\mathbf{A}$ matrix, we define $\mathbf{A}$ as the following matrix which has rank 2:

$$\mathbf{A} \equiv \boldsymbol{\alpha}\boldsymbol{\alpha}^T - \boldsymbol{\beta}\boldsymbol{\beta}^T \tag{84}$$

where the transposes of the column vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are defined as $\boldsymbol{\alpha}^T \equiv [\alpha_1, \ldots, \alpha_{80}]$ and $\boldsymbol{\beta}^T \equiv [\beta_1, \ldots, \beta_{80}]$. Thus we have reduced the number of unknown parameters in $\mathbf{A}$ from $(80+1) \times 80/2$ to 2×80.

With $\mathbf{A}$ defined by (84), the system of inverse demand *share* equations (46) becomes the following system of estimating equations:

$$s_{tn} = q_{tn}[\alpha_n \boldsymbol{\alpha} \cdot \boldsymbol{q}^t - \beta_n \boldsymbol{\beta} \cdot \boldsymbol{q}^t]/[(\boldsymbol{\alpha} \cdot \boldsymbol{q}^t)^2 - (\boldsymbol{\beta} \cdot \boldsymbol{q}^t)^2] + e_{tn};\ t = 1, \ldots, 144; n = 1, \ldots, 80. \tag{85}$$

Equations (85) are valid even when there are missing products because when product n is missing in period t, $s_{tn} = q_{tn} = 0$.

The utility function, $f(\boldsymbol{q}, \boldsymbol{\alpha}, \boldsymbol{\beta})$ is defined as follows:

$$f(\boldsymbol{q}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \equiv [\boldsymbol{q}^T(\boldsymbol{\alpha}\boldsymbol{\alpha}^T - \boldsymbol{\beta}\boldsymbol{\beta}^T)\boldsymbol{q}]^{1/2} = [(\boldsymbol{\alpha} \cdot \boldsymbol{q}^t)^2 - (\boldsymbol{\beta} \cdot \boldsymbol{q}^t)^2]^{1/2}. \tag{86}$$

Note that if $\boldsymbol{\beta} = \mathbf{0}_N$, then $f(\boldsymbol{q}, \boldsymbol{\alpha}, \mathbf{0}_N) = [(\boldsymbol{\alpha} \cdot \boldsymbol{q}^t)^2]^{1/2} = \boldsymbol{\alpha} \cdot \boldsymbol{q}^t = \sum_{n=1}^{N} \alpha_n q_n$; i.e., the utility function collapses down to the linear utility function that was studied in section 7. This is an important point because it implies that starting coefficients α_n and β_n for the nonlinear least squares minimization problem that is defined below can be set equal to the estimates of the α_n^* that result in the estimation of linear preferences with the starting coefficients for the β_n^* set equal to 0.

There are some tricky aspects to the new utility function as compared to the case of a linear utility function. We need to ensure that $(\boldsymbol{\alpha} \cdot \boldsymbol{q}^t)^2 - (\boldsymbol{\beta} \cdot \boldsymbol{q}^t)^2 > 0$ so that we can calculate the positive square root, $[(\boldsymbol{\alpha} \cdot \boldsymbol{q}^t)^2 - (\boldsymbol{\beta} \cdot \boldsymbol{q}^t)^2]^{1/2}$. We also need to set $\beta_n = 0$ if product n is available in only one period.*46 However, in our data set, all products are available for at least 14 periods. In order to identify all of the parameters, we impose our usual normalization so that our present model contains our linear utility function model as a special case:

$$\alpha_4 = 1 \tag{87}$$

Define the total sample consumption vector $\boldsymbol{q}^*$ as $\sum_{t=1}^{144} \boldsymbol{q}^t$. In order to prevent multicollinearity between the α_n and β_n parameters, we imposed the following normalization on the β_n parameters:

$$\boldsymbol{\beta} \cdot \boldsymbol{q}^* = 0.\text{*47} \tag{88}$$

Estimates for the α_n and β_n parameters are obtained by solving the following nonlinear least squares minimization problem subject to the normalizations (87) and (88):

$$\min_{\alpha,\beta} \sum_{t=1}^{T} \sum_{n=1}^{N} \{s_{tn} - q_{tn}[\alpha_n \boldsymbol{\alpha} \cdot \boldsymbol{q}^t - \beta_n \boldsymbol{\beta} \cdot \boldsymbol{q}^t]/[(\boldsymbol{\alpha} \cdot \boldsymbol{q}^t)^2 - (\boldsymbol{\beta} \cdot \boldsymbol{q}^t)^2]\}^2. \tag{89}$$

*46 If a product n appears in only one period in the sample of observations, then our KBF model will be able to estimate the parameter α_n but it will not be able to estimate the parameter β_n; see the Appendix in Diewert (2024)[26].

*47 This normalization also helps to ensure that $(\boldsymbol{\alpha} \cdot \boldsymbol{q}^t)^2 - (\boldsymbol{\beta} \cdot \boldsymbol{q}^t)^2 > 0$ so that we can define $f(\boldsymbol{q}^t)$ as the positive square root of $(\boldsymbol{\alpha} \cdot \boldsymbol{q}^t)^2 - (\boldsymbol{\beta} \cdot \boldsymbol{q}^t)^2$. This normalization ensures that our estimated KBF Prefecture price and quantity indexes are invariant to changes in the units of measurement. In our regression, used the constraint $\sum_{n=1}^{80} q_n^* \beta_n = 0$ to solve for $\beta_4 = -\sum_{n=1}^{3}[q_n^*/q_4^*]\beta_n - \sum_{n=5}^{80}[q_n^*/q_4^*]\beta_n$.

Taking into account the normalizations (87) and (88), there are 158 free parameters to estimate. The starting coefficient values for the α_n were the final coefficient estimates for the linear utility function model discussed in section 7 and the starting coefficient values for the β_n were set equal to 0.00001 or -0.00001. As a check on our code, the starting log likelihood for the model defined by (89) was equal to the final log likelihood for the linear model defined by (50) in section 7. Shazam took 393 iterations and 26 minutes to converge to a solution. The gain in log likelihood was 4289.18 points for adding 79 new β_n parameters. The R^2 for the new model was 0.9983, an increase over the R^2 for the linear model in section 7 ($R^2 = 0.9964$) and for the CES utility function model in the previous section ($R^2 = 0.9867$). The sum of absolute errors for the present model was 5.630 and for the linear model and the CES model, the sums were 7.9178 and 7.2587 respectively. Thus the present KBF model fits the data significantly better than previous models.

Once the solution $(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)$ to the nonlinear least squares minimization problem (89) has been obtained, the preliminary period t aggregate quantity and price levels, Q^{t^*} and P^{t^*}, are defined as follows:

$$Q^{t^*} \equiv f(\boldsymbol{q}^t, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) = [(\boldsymbol{\alpha}^* \cdot \boldsymbol{q}^t)^2 - (\boldsymbol{\beta}^* \cdot \boldsymbol{q}^t)^2]^{1/2}; P^{t^*} \equiv e^t / Q^{t^*}; \quad t = 1, \ldots, 144. \tag{90}$$

Normalize the sequence price levels P^{t^*} into the series P^t_{KBF} which is such that the normalized sequence of price levels equals 1 for $t = 25$ (month 1 for Tokyo):

$$P^t_{\mathrm{KBF}} \equiv P^{t^*} / P^{25^*}; \qquad t = 1, \ldots, 144. \tag{91}$$

Finally define the *KBF utility function price levels for regions 1-6* as $P^{r,m}_{\mathrm{KBF}}$ using P^t_{KBF} defined by (91) and an appropriate modification of definitions (64). The KBF price indexes, P^t_{KBF}, are plotted on Chart 6 below and the regional KBF indexes $P^{1,m}_{\mathrm{KBF}} - P^{6,m}_{\mathrm{KBF}}$ are listed in Table 8 in the Appendix.

Before we discuss Chart 6, we note that it is of interest to calculate the reservation prices that the estimated KBF utility function generates. With the solution $(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)$ to (89) in hand, we can calculate Hicksian reservation prices p^*_{tn} for the products n that were *not* present in period t using equations (83) for our BF functional form for products n that are not available in region-period t:

$$p^*_{tn} \equiv e^t f_n(\boldsymbol{q}^t, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) / f(\boldsymbol{q}^t, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*); \qquad t = 1, \ldots, T; n \notin S(t). \tag{92}$$

The average reservation price for our estimated KBF utility function turned out to equal 0.43135 while the average predicted price for products that were present in each period was equal to 0.32779. Thus on average, reservation prices were 0.43135/0.32779 = 1.316 or 31.6 percent higher than predicted prices.[*48] Note that the CES model generates infinite reservation prices which is a problem with the CES model.[*49]

The N by N matrix of second order partial derivatives of $f(\boldsymbol{q}, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)$ evaluated at $\boldsymbol{q} = \boldsymbol{q}^t$ is denoted by $\nabla^2 f(\boldsymbol{q}^t, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)$ and it is called the *period t inverse substitution matrix*. For a general linearly homogeneous and concave utility $f(\boldsymbol{q})$, it must be a negative semidefinite matrix that satisfies the restrictions $\nabla^2 f(\boldsymbol{q}^t)\boldsymbol{q}^t = \mathbf{0}_N$. Thus the rank of $\nabla^2 f(\boldsymbol{q}^t)$ is at most

[*48] The ratio of the average of observed prices to the average of predicted prices was 1.0038.

[*49] We also calculated the ratio of reservation prices to the average of predicted prices for the linear utility function model that was discussed in section 7. This ratio was 0.40690/0.32549 = 1.2501 so reservation prices for the linear utility function model were on average 25.0 percent higher. This implies that purchasers for the most part bought products when they were offered at discounted prices.

$N-1$. For our particular functional form for $f(\boldsymbol{q}, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)$ defined by (90), the period t inverse substitution matrix is defined as follows:

$$\nabla^2 f(\boldsymbol{q}^t, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) \equiv -[f(\boldsymbol{q}^t, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)]^{-3}[\boldsymbol{\alpha}^*(\boldsymbol{\beta}^* \cdot \boldsymbol{q}^t) - \boldsymbol{\beta}^*(\boldsymbol{\alpha}^* \cdot \boldsymbol{q}^t)][\boldsymbol{\alpha}^*(\boldsymbol{\beta}^* \cdot \boldsymbol{q}^t) - \boldsymbol{\beta}^*(\boldsymbol{\alpha}^* \cdot \boldsymbol{q}^t)]^T. \quad (93)$$

If $\boldsymbol{\beta}^* = \mathbf{0}_N$, then $\nabla^2 f(\boldsymbol{q}^t, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) = \mathbf{0}_N \mathbf{0}_N^T$ which is an N by N matrix of zeros. If $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$ are both nonzero vectors and $\boldsymbol{\alpha}^* \neq \boldsymbol{\beta}^*$, then the period t substitution matrix defined by (93) will have rank equal to one. Diewert and Wales (1988)[32] called a functional form for a cost function defined over N products a *semiflexible functional form of rank k* if its matrix of second order partial derivatives had rank k. Using this terminology, our $f(\boldsymbol{q}, \boldsymbol{\alpha}, \boldsymbol{\beta})$ defined by (90) is a *semiflexible functional form of rank 1.*[*50]

The KBF price indexes, P_{KBF}^t, are plotted on Chart 6 below along with five other price indexes that were described in previous sections. These indexes are the Linear Utility Function indexes P_{LU}^t, the CES price indexes P_{CES}^t that were defined by deflating region-period expenditures e^t by the estimated CES utility levels Q_{CES}^t, the Geary Khamis price indexes P_{GK}^t, the Implicit Weighted Time Product Dummy indexes P_{IWTPD}^t and the GEKS indexes P_{GEKS}^t.

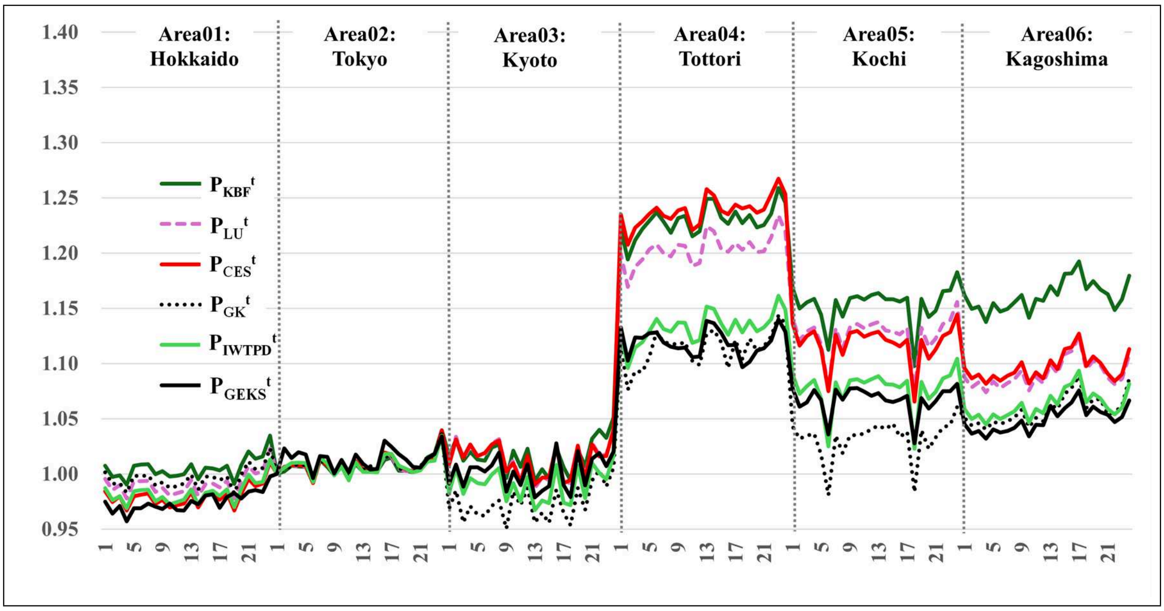

Chart 6 KBF Price Indexes and Other Indexes for Six Japanese Prefectures

The six series of price indexes are quite close to each other for the Tokyo Prefecture months (t = 25-48) and somewhat close for Hokkaido (t = 1-24) and Kyoto (t = 49-72) but very different for the three smaller population Prefectures, Tottori (t = 73-96), Kochi (t = 97-120) and Kagoshima (t = 121-144). It can be seen that the KBF indexes, P_{KBF}^t, are clearly higher for the Kochi and Kagoshima time periods and in general, are the highest price indexes. For periods 73-96, the KBF indexes are about 15 percentage points above the lowest index for most periods, the Geary Khamis indexes P_{GK}^t. The CES indexes P_{CES}^t and the Linear Utility function indexes P_{LU}^t are in general quite close and are the second highest indexes. The

[*50] If $f(q^t, \alpha^*, \beta^*) > 0$, then $\left[f(q^t, \alpha^*, \beta^*)\right]^{-3} \equiv k > 0$ and $A = k\gamma\gamma^T$ where γ is the N dimensional column vector $\alpha^*\left(\beta^* \cdot q^t\right) - \beta^*\left(\alpha^* \cdot q^t\right)$. Thus $z^T A z = z^T k\gamma\gamma^T z = k\left(z^T\gamma\right)^2 > 0$ for N dimensional vectors z. Thus A is a negative semidefinite symmetric matrix.

Implicit Weighted Time Product Dummy indexes P^t_{IWTPD} are well below P^t_{KBF}, P^t_{CES} and P^t_{LU} but above P^t_{GEKS} and P^t_{GK} for the smaller Prefectures. The fact that the six indexes are so dispersed is cause for concern: it means that the choice of method matters a lot.

The fact that the KBF indexes are well above the P^t_{LU}, P^t_{GK} and P^t_{IWTPD} indexes is not surprising if we take the economic approach to index number theory because the latter methods are based on their consistency with linear preferences which will tend to overstate the degree of product substitutability and hence bias the price indexes up and the corresponding quantity indexes down. However, it is surprising that the GEKS indexes were not higher since in theory, they are exact for a wide variety of preferences. Of course, the theoretical superiority of the GEKS indexes was established in the context of no missing products and the absence of discounted prices, which can lead to stockpiling and downward chain drift.

What is clear from Chart 6 is that constant quality rice prices in the low population Prefectures are much higher than prices in the higher population Prefectures. If this result persists over many products, it could mean that national Consumer Price Indexes constructed using current methods that do not adjust prices for product availability have a downward bias and hence national real consumption may have an upward bias.

10 Conclusion

Here are our tentative conclusions that we draw on from the above analysis:

- Indexes which do not weight prices by their economic importance can be unreliable. Their use should be avoided if possible.
- In the context of forming inter-regional price indexes where choice sets are very different, unit value and average price indexes can be very unreliable.
- GEKS indexes can also be unreliable in our context because matching between regions can be low if choice sets differ a lot across regions. GEKS (and CCDI) indexes rely on matched models and if there are few interregional matched products, these multilateral indexes can be inaccurate. Also, these indexes cannot measure the benefits of a larger choice set.
- Weighted and Unweighted Time Product Dummy Indexes also cannot completely measure the benefits of larger choice sets.
- The difference between our Linear Utility price index and the counterpart KBF Rank 1 Substitution Matrix Utility price index can measure the benefits of increased choice sets and the use of a more flexible functional form. But it is not easy to do the econometric estimation for the KBF model if there are a large number of products in scope.
- The estimation of CES preferences is also problematic: different econometric specifications can generate very different estimates for the elasticity of substitution and hence can generate very different estimates for the gains from increased variety using the Feenstra methodology.
- On the other hand, Weighted Time Product Dummy price indexes can be estimated when there are numerous products but these indexes cannot measure accurately the benefits of larger choice sets.
- A robust method for dealing with rapid product turnover, quality adjustment and the chain drift problem is still to be found.

References

[1] Arrow, K.J., H.B. Chenery, B.S. Minhas and R.M. Solow (1961), "Capital-Labor Substitution and Economic Efficiency", *Review of Economics and Statistics* 63, 225-250.

[2] Australian Bureau of Statistics (2016), "Making Greater Use of Transactions Data to Compile the Consumer Price Index", Information Paper 6401.0.60.003, November 29, Canberra: ABS.

[3] Balk, B.M. (1980), "A Method for Constructing Price Indices for Seasonal Commodities", *Journal of the Royal Statistical Society*, Series A 143, 68-75.

[4] Balk, B.M. (1981), "A Simple Method for Constructing Price Indices for Seasonal Commodities", *Statistische Hefte* 22 (1), 1–8.

[5] Balk, B.M. (1996), "A Comparison of Ten Methods for Multilateral International Price and Volume Comparisons", *Journal of Official Statistics* 12, 199-222.

[6] Balk, B.M. (1999), "On Curing the CPI's Substitution and New Goods Bias", Paper presented at the Fifth Meeting of the Ottawa Group, Reykjavik, Iceland, 1999.

[7] Balk, B.M. (2008), *Price and Quantity Index Numbers*, New York: Cambridge University Press.

[8] Balk, B.M. (2024), "Extension Methods for Multilateral Price Indices and the Return of Chain Drift", paper presented at the UNSW-ESCoE Conference on Economic Measurement 2024, November 25, University of New South Wales, Sydney Australia.

[9] Chessa, A.G. (2016), "A New Methodology for Processing Scanner Data in the Dutch CPI", *EURONA* 1, 49–69.

[10] Chessa, A.G. (2021), "Extension of Multilateral Index Series over Time: Analysis and Comparison of Methods, United Nations ECE/ILO Meeting of the Group of Experts on Consumer Price Indices, Geneva, 2-10 June 2021.

[11] Court, A.T. (1939), "Hedonic Price Indexes with Automotive Examples", pp. 99-117 in *The Dynamics of Automobile Demand*, New York: General Motors Corporation.

[12] de Haan, J. (2008), "Reducing Drift in Chained Superlative Price Indexes for Highly Disaggregated Data", paper presented at the Economic Measurement Workshop, Centre for Applied Economic Research, University of New South Wales, December 10.

[13] de Haan, J. (2015), "Rolling Year Time Dummy Indexes and the Choice of Splicing Method", paper presented at the 14th Meeting of the Ottawa Group, Tokyo, Japan. website: `https://stats.unece.org/ottawagroup/meeting/14`

[14] de Haan, J. and H. van der Grient (2011), "Eliminating Chain Drift in Price Indexes Based on Scanner Data", *Journal of Econometrics* 161, 36-46.

[15] De Haan, J. and F. Krsinich (2018), "Time Dummy Hedonic and Quality Adjusted Unit Value Indexes: Do They Really Differ?", *Review of Income and Wealth*, Series 64:4, 757-776.

[16] De Haan, J. and F. Krsinich (2024), "Product Churn and the GEKS-Törnqvist Price Index: The Feenstra Adjustment", paper presented at the 18th Meeting of the Ottawa Group, Ottawa, Canada. website: `https://stats.unece.org/ottawagroup/meeting/18`

[17] Diewert, W.E. (1974), "Applications of Duality Theory," pp. 106-171 in M.D. Intriligator and D.A. Kendrick (ed.), *Frontiers of Quantitative Economics*, Vol. II, Amsterdam: North-Holland.

[18] Diewert, W.E. (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics* 4, 114-145.

[19] Diewert, W.E. (1999), "Axiomatic and Economic Approaches to International Compar-

isons", 13-87 in *International and Interarea Comparisons of Income, Output and Prices*, A. Heston and R.E. Lipsey (eds.), Studies in Income and Wealth, Volume 61, Chicago: The University of Chicago Press.

[20] Diewert, W.E. (2004), "On the Stochastic Approach to Linking the Regions in the ICP", Discussion Paper 04-16, Department of Economics, The University of British Columbia, Vancouver, Canada.

[21] Diewert, W.E. (2005), "Weighted Country Product Dummy Variable Regressions and Index Number Formulae", *Review of Income and Wealth* 51, 561-570.

[22] Diewert, W.E. and R.C. Feenstra (2017), "Estimating the Benefits and Costs of New and Disappearing Products", Discussion Paper 17-10, Vancouver School of Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1L4.

[23] Diewert, W.E. (2020a), "The Economic Approach to Index Number Theory", Discussion Paper 20-05, Vancouver School of Economics, The University of British Columbia, Vancouver, Canada V6T 1L4.

[24] Diewert, W.E. (2020b), "Quality Adjustment Methods", Discussion Paper 20-08, Vancouver School of Economics, The University of British Columbia, Vancouver, Canada V6T 1L4.

[25] Diewert, W.E. (2023), "Scanner Data, Elementary Price Indexes and the Chain Drift Problem", pp. 445-606 in *Advances in Economic Measurement*, D. Chotikapanich, A.N. Rambaldi and N. Rhode (eds.), Singapore: Palgrave Macmillan.

[26] Diewert, W.E. (2024), "The Ottawa Group After 30 Years", Discussion Paper 24-02, Vancouver School of Economics, University of British Columbia, 6000 Iona Drive, Vancouver B.C., Canada V6T 1L4.

[27] Diewert, W.E. and R.C. Feenstra (2017), "Estimating the Benefits and Costs of New and Disappearing Products", Discussion Paper 17-10, Vancouver School of Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1L4.

[28] Diewert, W.E. and R.C. Feenstra (2022), "Estimating the Benefits of New Products", pp. 437-473 in *Big Data For the Twenty-First Century Economic Statistics*, K.G. Abraham, R.S. Jarmin, B.C. Moyer and M.D. Shapiro (eds.), Chicago: University of Chicago Press.

[29] Diewert, W.E. and K.J. Fox (2017) "Substitution Bias in Multilateral Methods for CPI Construction Using Scanner Data", Discussion Paper 17-02, Vancouver School of Economics, The University of British Columbia, Vancouver, Canada, V6T 1L4.

[30] Diewert, W.E. and K.J. Fox (2021) "Substitution Bias in Multilateral Methods for CPI Construction Using Scanner Data," *Journal of Business and Economic Statistics* 40:1, 355-369.

[31] Diewert, W.E. and C. Shimizu (2024), "Scanner Data, Product Churn and Quality Adjustment", 18th Ottawa Group Meeting, Ottawa, May 13.

[32] Diewert, W.E. and T.J. Wales (1988), "A Normalized Quadratic Semiflexible Functional Form", *Journal of Econometrics* 37, 327-42.

[33] Diewert, W.E. and V. Zelenyuk (2024), "On the Problem of the Purchasing Power of Money by A. A. Konüs and S. S. Byushgens: Translation and Commentary", *Journal of Productivity Analysis*, `https://doi.org/10.1007/s11123-023-00696-x`

[34] Eltetö, Ö., and Köves, P. (1964), "On a Problem of Index Number Computation Relating to International Comparisons", (in Hungarian), *Statisztikai Szemle* 42, 507-518.

[35] Feenstra, R.C. (1994), "New Product Varieties and the Measurement of International Prices", *American Economic Review* 84:1, 157-177.

[36] Fisher, I. (1922), *The Making of Index Numbers*, Houghton-Mifflin, Boston.

[37] Fox, K.J., P. Levell and M. O'Connell (2024), "Inflation Measurement with High Frequency Data", 18th Ottawa Group Meeting, Ottawa, May 13. `https://stats.unece.org/ottawagroup/meeting/18`

[38] Frisch, R. (1936), "Annual Survey of General Economic Theory: The Problem of Index Numbers", *Econometrica* 4, 1-39.

[39] Frobenius, G. (1912), "Über Matrizen aus nicht negative Elementen", *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, 471-476.

[40] Gantmacher, F. (1959), *The Theory of Matrices, Volume 2*, New York: Interscience Publishers.

[41] Geary, R.G. (1958), "A Note on Comparisons of Exchange Rates and Purchasing Power between Countries", *Journal of the Royal Statistical Society Series A* 121, 97-99.

[42] Gini, C. (1924), "Quelques considérations au sujet de la construction des nombres indices des prix et des questions analogues", *Metron* 4:1, 3-162.

[43] Gini, C. (1931), "On the Circular Test of Index Numbers", *Metron* 9:9, 3-24.

[44] Hardy, G.H., J.E. Littlewood and G. Polyá (1934), *Inequalities*, Cambridge: Cambridge University Press.

[45] Hausman, J.A. (1996), "Valuation of New Goods under Perfect and Imperfect Competition", pp. 20 -236 in *The Economics of New Goods*, T.F. Bresnahan and R.J. Gordon (eds.), Chicago: University of Chicago Press.

[46] Hausman, J.A. (1999), "Cellular Telephone, New Products and the CPI", *Journal of Business and Economic Statistics* 17:2, 188-194.

[47] Hicks, J.R. (1940), "The Valuation of the Social Income", *Economica* 7, 105-140.

[48] Hill, R.J. (2001), "Measuring Inflation and Growth Using Spanning Trees", *International Economic Review* 42, 167-185.

[49] Hill, R.J. (2004), "Constructing Price Indexes Across Space and Time: The Case of the European Union", *American Economic Review* 94, 1379-1410.

[50] Hill, T.P. (1988), "Recent Developments in Index Number Theory and Practice", *OECD Economic Studies* 10, 123-148.

[51] Ivancic, L., W.E. Diewert and K.J. Fox (2009), "Scanner Data, Time Aggregation and the Construction of Price Indexes", Discussion Paper 09-09, Department of Economics, University of British Columbia, Vancouver, Canada.

[52] Ivancic, L., W.E. Diewert and K.J. Fox (2011), "Scanner Data, Time Aggregation and the Construction of Price Indexes", *Journal of Econometrics* 161, 24-35.

[53] Khamis, S.H. (1970), "Properties and Conditions for the Existence of a New Type of Index Number", *Sankhya* B 32, 81-98.

[54] Konüs, A.A. and S.S. Byushgens (1926), "K probleme pokupatelnoi sily deneg", *Voprosy Konyunktury* 2(1), 151-172.

[55] Krsinich, F. (2016), "The FEWS Index: Fixed Effects with a Window Splice", *Journal of Official Statistics* 32, 375-404.

[56] Melser, D. (2006), "Accounting for the Effects of New and Disappearing Goods using Scanner Data", *Review of Income and Wealth* Series 52:4, 547-568.

[57] Melser, D. and M. Webster (2021), "Multilateral Methods, Substitution Bias and Chain Drift: Some Empirical Examples", *Review of Income and Wealth* 67:3, 759-785.

[58] Persons, W.M. (1921), "Fisher's Formula for Index Numbers", *Review of Economics and Statistics* 3:5, 103-113.

[59] Persons, W.M. (1928), "The Effect of Correlation Between Weights and Relatives in the Construction of Index Numbers", *Review of Economics and Statistics* 10:2, 80-107.

[60] Rao, D.S. Prasada (1995), "On the Equivalence of the Generalized Country-Product-Dummy (CPD) Method and the Rao-System for Multilateral Comparisons", Working Paper No. 5, Centre for International Comparisons, University of Pennsylvania, Philadelphia.

[61] Rao, D.S. Prasada (2005), "On the Equivalence of the Weighted Country Product Dummy (CPD) Method and the Rao System for Multilateral Price Comparisons", *Review of In-*

come and Wealth 51:4, 571-580.

[62] Samuelson, P.A. and S. Swamy (1974), "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis", *American Economic Review* 64, 566-593.

[63] Summers, R. (1973), "International Comparisons with Incomplete Data", *Review of Income and Wealth* 29:1, 1-16.

[64] Szulc, B.J. (1964), "Indices for Multiregional Comparisons", (in Polish), *Przeglad Statystyczny* 3, 239-254.

[65] Szulc, B.J. (1983), "Linking Price Index Numbers," pp. 537-566 in *Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada.

[66] Szulc, B.J. (1987), "Price Indexes below the Basic Aggregation Level", *Bulletin of Labour Statistics* 2, 9-16.

[67] Walsh, C.M. (1901), *The Measurement of General Exchange Value*, New York: Macmillan and Co.

[68] Walsh, C.M. (1921), "*Discussion*", Journal of the American Statistical Association 17, 537-544.

[69] White, K.J. (2004), *Shazam: User's Reference Manual, Version 10*, Vancouver, Canada: Northwest Econometrics Ltd.

[70] Zhang, L.-C., I. Johansen and R. Nygaard (2019), "Tests for Price Indices in a Dynamic Item Universe", *Journal of Official Statistics* 35:3, 683–697.

Online Appendix: Index Tables

Table 1: Average Price, Unit Value, Fixed Base Fisher, GEKS and Fisher-Mizobuchi Price Indexes for Six Japanese Prefectures

m	$P_{\mathrm{AV}}^{1,m}$	$P_{\mathrm{UV}}^{1,m}$	$P_{\mathrm{F}}^{1,m}$	$P_{\mathrm{GEKS}}^{1,m}$	$P_{\mathrm{FM}}^{1,m}$	$P_{\mathrm{AV}}^{2,m}$	$P_{\mathrm{UV}}^{2,m}$	$P_{\mathrm{F}}^{2,m}$	$P_{\mathrm{GEKS}}^{2,m}$	$P_{\mathrm{FM}}^{2,m}$
1	0.99049	1.00350	0.99526	0.97500	0.99526	1.00000	1.00000	1.00000	1.00000	1.00000
2	0.98259	0.99258	0.98266	0.96424	0.97550	1.00899	1.00439	1.00461	1.02316	1.00461
3	0.99068	0.99363	0.98611	0.97102	0.97762	1.00494	1.00693	1.00593	1.01437	1.00497
4	0.98601	0.98468	0.98110	0.95714	0.97646	1.00740	1.00868	1.01083	1.01978	1.01026
5	0.99042	1.00051	0.98845	0.96913	0.98293	1.00604	1.00761	1.00868	1.01753	1.00772
6	0.99419	1.00100	0.98868	0.96904	0.98669	0.99223	0.99285	0.99049	0.99626	0.99135
7	0.99575	1.00147	0.99398	0.97335	0.99179	1.00091	1.01294	1.01227	1.01615	1.01169
8	0.99210	0.99253	0.99201	0.97039	0.98714	1.00130	1.00900	1.00866	1.01546	1.00796
9	0.99413	0.99520	0.98584	0.96830	0.98256	0.99619	1.00154	0.99959	1.00047	1.00001
10	0.99481	0.99026	0.99410	0.97310	0.98391	1.00181	1.00840	1.00645	1.01261	1.00726
11	0.98246	0.99157	0.98910	0.96743	0.98669	0.99157	0.99921	0.99686	1.00156	0.99888
12	0.99495	0.99398	0.98854	0.96697	0.97798	1.00523	1.00893	1.00918	1.01749	1.01333
13	0.99301	1.00192	0.99511	0.97576	0.98828	0.99640	1.00387	1.00063	1.00936	1.00355
14	0.99231	0.98882	0.98849	0.97240	0.98214	0.99707	1.00681	1.00350	1.00397	1.00663
15	0.99883	1.00030	0.99048	0.98062	0.98726	1.00078	1.00442	1.00354	1.00486	1.00416
16	0.99118	0.99984	0.98570	0.98151	0.98237	1.01604	1.01503	1.01818	1.03016	1.02136
17	0.99513	0.99704	0.98193	0.96952	0.97335	1.01254	1.01748	1.01885	1.02465	1.01976
18	0.99901	1.00008	0.98944	0.97917	0.97638	1.00089	1.00454	1.00723	1.01816	1.00920
19	0.99919	0.98492	0.99479	0.98261	0.97933	1.00522	1.00390	1.00495	1.01322	1.00667
20	1.00004	1.00128	0.99007	0.97796	0.98912	0.99627	1.00276	1.00436	1.00637	1.00424
21	1.00946	1.01357	0.99856	0.98393	1.00056	1.00862	1.00647	1.00550	1.00561	1.00815
22	1.00513	1.00672	1.00126	0.98552	1.00597	1.02060	1.01549	1.01767	1.01433	1.02012
23	1.01500	1.00752	1.00217	0.98374	1.01286	1.02260	1.01770	1.01867	1.01817	1.02121
24	1.02921	1.02561	1.01851	0.99767	1.02833	1.02641	1.03643	1.03694	1.03397	1.03884
Mean	0.99650	0.99869	0.99176	0.99176	0.98794	1.00500	1.00810	1.00810	1.01320	1.00920

Table 1: Continued

m	$P_{\mathrm{AV}}^{3,m}$	$P_{\mathrm{UV}}^{3,m}$	$P_{\mathrm{F}}^{3,m}$	$P_{\mathrm{GEKS}}^{3,m}$	$P_{\mathrm{FM}}^{3,m}$	$P_{\mathrm{AV}}^{4,m}$	$P_{\mathrm{UV}}^{4,m}$	$P_{\mathrm{F}}^{4,m}$	$P_{\mathrm{GEKS}}^{4,m}$	$P_{\mathrm{FM}}^{4,m}$
1	1.01934	0.97410	0.97619	0.99154	0.97619	1.04608	1.11598	1.15919	1.13288	1.15919
2	1.01612	0.99042	0.98083	1.00849	0.97092	1.03793	1.07604	1.12359	1.10251	1.06255
3	1.01512	0.96182	0.95948	0.98863	0.96794	1.05509	1.09225	1.14505	1.12347	1.11954
4	1.01927	0.97655	0.97299	1.00624	0.96736	1.05028	1.09233	1.14682	1.12336	1.12217
5	1.00825	0.96831	0.98283	1.00619	0.98205	1.05669	1.10540	1.14702	1.12736	1.11867
6	1.01246	0.96697	0.98398	1.00233	0.99284	1.05896	1.12923	1.14884	1.12811	1.11673
7	1.01444	0.97641	0.98543	1.00778	0.99399	1.05489	1.11961	1.13821	1.11902	1.11782
8	1.02200	0.98104	0.99593	1.01935	0.98355	1.04975	1.11811	1.13846	1.11527	1.11871
9	1.01845	0.95506	0.98397	0.98407	0.97209	1.05576	1.11692	1.12784	1.11379	1.12478
10	1.02414	0.98825	1.01598	1.00236	1.00526	1.05417	1.11927	1.12933	1.11452	1.11038
11	1.02445	0.97557	1.00624	0.99010	0.99318	1.04571	1.10185	1.11917	1.10588	1.10575
12	1.02979	0.99138	1.02312	1.00877	1.00495	1.05150	1.09892	1.12118	1.10652	1.09936
13	1.02394	0.96040	1.00002	0.97895	0.98253	1.05717	1.12679	1.15844	1.13877	1.13341
14	1.02563	0.96848	1.00627	0.98489	0.99759	1.05804	1.13056	1.15697	1.13647	1.15639
15	1.03210	0.95930	1.00155	0.98900	0.98907	1.05292	1.11938	1.14791	1.12779	1.14435
16	1.03486	0.98971	1.02654	1.02785	1.00187	1.06243	1.09563	1.13328	1.11648	1.08074
17	1.01663	0.96836	1.00265	0.98993	0.98502	1.05244	1.12107	1.12973	1.11683	1.10189
18	1.00236	0.95702	0.99375	0.97926	0.98066	1.04248	1.10406	1.10474	1.09659	1.06290
19	1.02109	0.99091	1.01947	1.02097	0.99865	1.04022	1.12231	1.11382	1.10170	1.09661
20	1.01832	0.97117	1.00604	0.98957	0.99528	1.05573	1.11158	1.12771	1.11176	1.10982
21	1.03252	0.99795	1.02113	1.01395	1.01237	1.04539	1.11582	1.13018	1.11370	1.13549
22	1.03873	1.00660	1.02972	1.01908	1.03233	1.08529	1.12443	1.14214	1.12096	1.17330
23	1.03364	0.99216	1.01863	1.00710	1.01208	1.09551	1.14331	1.15871	1.14018	1.17694
24	1.03842	1.00951	1.04238	1.02048	1.03459	1.07887	1.13597	1.15037	1.12875	1.16649
Mean	1.02260	0.97823	1.00150	1.00150	0.99301	1.05600	1.11400	1.13740	1.11930	1.12140

Table 1: Concluded

m	$P_{\mathrm{AV}}^{5,m}$	$P_{\mathrm{UV}}^{5,m}$	$P_{\mathrm{F}}^{5,m}$	$P_{\mathrm{GEKS}}^{5,m}$	$P_{\mathrm{FM}}^{5,m}$	$P_{\mathrm{AV}}^{6,m}$	$P_{\mathrm{UV}}^{6,m}$	$P_{\mathrm{F}}^{6,m}$	$P_{\mathrm{GEKS}}^{6,m}$	$P_{\mathrm{FM}}^{6,m}$
1	1.06652	1.04492	1.04876	1.07757	1.04876	1.03129	1.05067	1.03408	1.04529	1.03408
2	1.05936	1.03434	1.04207	1.06086	1.03185	1.02643	1.04396	1.02743	1.03659	1.02393
3	1.05589	1.03737	1.05032	1.06501	1.04728	1.02524	1.04595	1.03080	1.03896	1.02131
4	1.06991	1.03967	1.05423	1.07596	1.03984	1.02148	1.04022	1.02451	1.03209	1.01395
5	1.07991	1.01968	1.05157	1.06680	1.04232	1.02208	1.04693	1.02785	1.04048	1.02163
6	1.06415	0.98170	1.03549	1.03591	1.02844	1.02846	1.04474	1.02777	1.03773	1.03634
7	1.06932	1.03740	1.06212	1.07658	1.06206	1.02177	1.04730	1.02340	1.03882	1.02994
8	1.06223	1.02189	1.04619	1.06734	1.03857	1.03016	1.05105	1.03201	1.04180	1.02814
9	1.04202	1.03738	1.05522	1.07755	1.06008	1.02787	1.05808	1.03570	1.04845	1.03302
10	1.05545	1.03816	1.05316	1.07792	1.07241	1.01551	1.04204	1.02191	1.03405	1.00749
11	1.06418	1.03855	1.05562	1.07469	1.07998	1.02219	1.05025	1.03175	1.04468	1.02213
12	1.05838	1.04355	1.05570	1.07066	1.06767	1.02898	1.04730	1.03459	1.04437	1.02980
13	1.05990	1.04630	1.05869	1.07336	1.07464	1.03213	1.06255	1.04560	1.06166	1.03435
14	1.05182	1.04326	1.05845	1.06674	1.08806	1.03114	1.05856	1.03457	1.05228	1.01955
15	1.04670	1.04760	1.05809	1.06535	1.07651	1.02842	1.07193	1.04380	1.05950	1.03014
16	1.05904	1.03634	1.05318	1.06718	1.07230	1.03024	1.07753	1.04490	1.06475	1.03936
17	1.05907	1.04047	1.05589	1.07097	1.07602	1.03226	1.08820	1.05457	1.07584	1.04592
18	1.02718	0.98499	1.02481	1.02777	1.03352	1.03014	1.05922	1.03503	1.05349	1.02093
19	1.05185	1.04200	1.05650	1.06857	1.06103	1.03189	1.06731	1.03926	1.06101	1.02734
20	1.05027	1.02489	1.04638	1.05930	1.05393	1.02759	1.06720	1.03761	1.05603	1.02628
21	1.05983	1.03491	1.05884	1.06580	1.06215	1.02332	1.05373	1.03573	1.05370	1.03245
22	1.04348	1.04450	1.05588	1.07508	1.08021	1.02275	1.05530	1.02542	1.04692	1.02415
23	1.05981	1.04683	1.06322	1.07503	1.07325	1.03137	1.06153	1.03339	1.05163	1.03682
24	1.06355	1.06307	1.07558	1.08151	1.09159	1.03546	1.08621	1.04460	1.06678	1.04982
Mean	1.05750	1.03460	1.05320	1.06760	1.06090	1.02740	1.05740	1.03440	1.04950	1.02870

Table 2: Time Product Dummy, Direct and Indirect Weighted Time Product Dummy and GEKS Price Indexes for Six Japanese Prefectures

m	$P^{1,m}_{\text{TPD}}$	$P^{1,m}_{\text{WTPD}}$	$P^{1,m}_{\text{IWTPD}}$	$P^{1,m}_{\text{GEKS}}$	$P^{2,m}_{\text{TPD}}$	$P^{2,m}_{\text{WTPD}}$	$P^{2,m}_{\text{IWTPD}}$	$P^{2,m}_{\text{GEKS}}$
1	0.98384	0.98566	0.98724	0.97500	1.00000	1.00000	1.00000	1.00000
2	0.97515	0.97491	0.97642	0.96424	1.00834	1.00603	1.00589	1.02316
3	0.98356	0.97851	0.98004	0.97102	1.00503	1.00912	1.00982	1.01437
4	0.97861	0.96822	0.96907	0.95714	1.00690	1.01077	1.01040	1.01978
5	0.98368	0.98320	0.98458	0.96913	1.00605	1.00971	1.01004	1.01753
6	0.98702	0.98401	0.98533	0.96904	0.99278	0.99261	0.99289	0.99626
7	0.98914	0.98441	0.98575	0.97335	1.00085	1.01453	1.01516	1.01615
8	0.98481	0.97461	0.97536	0.97039	1.00258	1.00881	1.00922	1.01546
9	0.98642	0.97798	0.97903	0.96830	0.99754	0.99905	0.99871	1.00047
10	0.98519	0.97216	0.97255	0.97310	1.00388	1.00605	1.00697	1.01261
11	0.97208	0.97367	0.97467	0.96743	0.99414	0.99473	0.99439	1.00156
12	0.98569	0.97592	0.97690	0.96697	1.00748	1.00972	1.01061	1.01749
13	0.98331	0.98483	0.98617	0.97576	0.99885	1.00125	1.00192	1.00936
14	0.98302	0.97129	0.97226	0.97240	0.99923	1.00207	1.00189	1.00397
15	0.99069	0.98143	0.98288	0.98062	1.00352	1.00109	1.00160	1.00486
16	0.98136	0.98315	0.98469	0.98151	1.01719	1.01721	1.01802	1.03016
17	0.98590	0.97919	0.98035	0.96952	1.01458	1.01716	1.01837	1.02465
18	0.99066	0.98477	0.98622	0.97917	1.00366	1.00598	1.00760	1.01816
19	0.99096	0.96964	0.97008	0.98261	1.00763	1.00352	1.00456	1.01322
20	0.99178	0.98505	0.98624	0.97796	0.99881	1.00074	1.00152	1.00637
21	1.00115	0.99832	0.99952	0.98393	1.01065	1.00392	1.00342	1.00561
22	0.99445	0.99081	0.99190	0.98552	1.02238	1.01212	1.01128	1.01433
23	1.00507	0.99215	0.99299	0.98374	1.02606	1.01296	1.01210	1.01817
24	1.01923	1.01206	1.01305	0.99767	1.02987	1.03486	1.03465	1.03397
Mean	0.98803	0.98192	0.98305	0.97481	1.00660	1.00730	1.00750	1.01320

Table 2: Continued

m	$P^{3,m}_{\text{TPD}}$	$P^{3,m}_{\text{WTPD}}$	$P^{3,m}_{\text{IWTPD}}$	$P^{3,m}_{\text{GEKS}}$	$P^{4,m}_{\text{TPD}}$	$P^{4,m}_{\text{WTPD}}$	$P^{4,m}_{\text{IWTPD}}$	$P^{4,m}_{\text{GEKS}}$
1	1.01586	0.98452	0.98273	0.99154	1.03955	1.12970	1.12934	1.13288
2	1.01324	1.00546	1.00555	1.00849	1.02950	1.09676	1.09588	1.10251
3	1.01274	0.98173	0.98200	0.98863	1.04864	1.11498	1.11466	1.12347
4	1.01750	0.99569	0.99627	1.00624	1.04120	1.12049	1.11952	1.12336
5	1.01079	0.99195	0.99172	1.00619	1.05126	1.12929	1.12934	1.12736
6	1.01553	0.99045	0.99059	1.00233	1.05455	1.13951	1.14055	1.12811
7	1.01983	0.99869	0.99934	1.00778	1.04934	1.13154	1.13141	1.11902
8	1.02247	1.00510	1.00554	1.01935	1.04498	1.12806	1.12891	1.11527
9	1.01930	0.97501	0.97523	0.98407	1.05053	1.13704	1.13725	1.11379
10	1.02237	0.99370	0.99168	1.00236	1.04934	1.13621	1.13679	1.11452
11	1.02318	0.97752	0.97538	0.99010	1.04045	1.11990	1.11888	1.10588
12	1.02887	0.99873	0.99725	1.00877	1.04490	1.11992	1.12107	1.10652
13	1.02427	0.96897	0.96724	0.97895	1.05040	1.15146	1.15155	1.13877
14	1.02366	0.97848	0.97609	0.98489	1.05203	1.14961	1.14944	1.13647
15	1.03010	0.97445	0.97360	0.98900	1.04498	1.13583	1.13621	1.12779
16	1.03278	1.00711	1.00819	1.02785	1.05466	1.12861	1.12618	1.11648
17	1.01939	0.97591	0.97410	0.98993	1.04857	1.13827	1.13968	1.11683
18	1.00403	0.97152	0.97171	0.97926	1.03860	1.12828	1.12831	1.09659
19	1.02199	1.00645	1.00763	1.02097	1.03287	1.13927	1.13902	1.10170
20	1.01809	0.97946	0.97804	0.98957	1.05032	1.12950	1.12939	1.11176
21	1.03239	1.00999	1.01013	1.01395	1.03964	1.13262	1.13260	1.11370
22	1.03845	1.00468	1.00076	1.01908	1.06392	1.14023	1.14016	1.12096
23	1.03514	0.99688	0.99511	1.00710	1.07412	1.16129	1.16151	1.14018
24	1.04168	1.01788	1.01711	1.02048	1.05621	1.14874	1.14873	1.12875
Mean	1.02270	0.99126	0.99054	1.00150	1.04790	1.13280	1.13280	1.11930

Table 2: Concluded

m	$P^{5,m}_{\text{TPD}}$	$P^{5,m}_{\text{WTPD}}$	$P^{5,m}_{\text{IWTPD}}$	$P^{5,m}_{\text{GEKS}}$	$P^{6,m}_{\text{TPD}}$	$P^{6,m}_{\text{WTPD}}$	$P^{6,m}_{\text{IWTPD}}$	$P^{6,m}_{\text{GEKS}}$
1	1.06408	1.08857	1.09028	1.07757	1.03035	1.05742	1.05878	1.04529
2	1.05695	1.07121	1.07243	1.06086	1.02640	1.04918	1.05000	1.03659
3	1.05369	1.07811	1.07969	1.06501	1.02551	1.05217	1.05365	1.03896
4	1.06637	1.08383	1.08528	1.07596	1.02096	1.04311	1.04483	1.03209
5	1.07223	1.06766	1.06845	1.06680	1.02260	1.05293	1.05404	1.04048
6	1.06453	1.02718	1.02505	1.03591	1.02827	1.04899	1.04977	1.03773
7	1.07042	1.08183	1.08319	1.07658	1.02054	1.05324	1.05330	1.03882
8	1.06334	1.06619	1.06749	1.06734	1.02909	1.05606	1.05668	1.04180
9	1.03071	1.08483	1.08478	1.07755	1.03135	1.06376	1.06461	1.04845
10	1.05012	1.08459	1.08590	1.07792	1.01645	1.04504	1.04674	1.03405
11	1.06311	1.08145	1.08280	1.07469	1.02652	1.05813	1.05896	1.04468
12	1.05821	1.08468	1.08600	1.07066	1.03350	1.05414	1.05501	1.04437
13	1.05995	1.08708	1.08854	1.07336	1.03448	1.06984	1.07073	1.06166
14	1.05177	1.08008	1.08136	1.06674	1.03350	1.06265	1.06311	1.05228
15	1.04850	1.07958	1.08096	1.06535	1.03289	1.07828	1.07885	1.05950
16	1.05645	1.07732	1.07845	1.06718	1.03165	1.08190	1.08164	1.06475
17	1.06305	1.08330	1.08436	1.07097	1.03298	1.09416	1.09363	1.07584
18	1.02633	1.02592	1.02303	1.02777	1.03156	1.06596	1.06483	1.05349
19	1.05669	1.08234	1.08354	1.06857	1.03291	1.07357	1.07275	1.06101
20	1.05446	1.06711	1.06800	1.05930	1.02958	1.06856	1.06853	1.05603
21	1.06436	1.07355	1.07423	1.06580	1.02355	1.06155	1.05895	1.05370
22	1.03943	1.08556	1.08668	1.07508	1.02332	1.05423	1.05401	1.04692
23	1.06145	1.08802	1.08932	1.07503	1.03214	1.06178	1.06003	1.05163
24	1.06475	1.10446	1.10447	1.08151	1.03630	1.08254	1.08153	1.06678
Mean	1.05670	1.07640	1.07730	1.06760	1.02860	1.06200	1.06230	1.04950

Table 3: Geary Khamis Price Indexes for Six Japanese Prefectures

m	$P^{1,m}_{\text{GK}}$	$P^{2,m}_{\text{GK}}$	$P^{3,m}_{\text{GK}}$	$P^{4,m}_{\text{GK}}$	$P^{5,m}_{\text{GK}}$	$P^{6,m}_{\text{GK}}$
1	1.00167	1.00000	0.96884	1.11595	1.04277	1.05079
2	0.99091	1.00456	0.98529	1.07639	1.03203	1.04393
3	0.99178	1.00691	0.95626	1.09304	1.03466	1.04607
4	0.98288	1.00875	0.97090	1.09236	1.03641	1.04049
5	0.99835	1.00773	0.96342	1.10610	1.01761	1.04729
6	0.99822	0.99292	0.96200	1.12884	0.98085	1.04515
7	0.99874	1.01286	0.97116	1.11933	1.03481	1.04771
8	0.98967	1.00907	0.97595	1.11817	1.01919	1.05133
9	0.99257	1.00163	0.95054	1.11721	1.03482	1.05839
10	0.98747	1.00859	0.98423	1.11934	1.03543	1.04236
11	0.98881	0.99949	0.97185	1.10230	1.03638	1.05079
12	0.99133	1.00918	0.98719	1.09885	1.04087	1.04785
13	0.99931	1.00418	0.95627	1.12727	1.04396	1.06312
14	0.98600	1.00704	0.96476	1.13038	1.04064	1.05905
15	0.99766	1.00465	0.95564	1.11955	1.04547	1.07243
16	0.99742	1.01534	0.98588	1.09615	1.03389	1.07801
17	0.99472	1.01770	0.96484	1.12151	1.03818	1.08865
18	0.99779	1.00482	0.95356	1.10441	0.98374	1.05962
19	0.98227	1.00411	0.98728	1.12259	1.03957	1.06779
20	0.99906	1.00292	0.96774	1.11165	1.02293	1.06758
21	1.01131	1.00667	0.99438	1.11570	1.03302	1.05388
22	1.00438	1.01564	1.00319	1.12475	1.04229	1.05598
23	1.00533	1.01787	0.98860	1.14369	1.04456	1.06184
24	1.02337	1.03665	1.00591	1.13626	1.06057	1.08647
Mean	0.99629	0.99629	0.97399	1.11420	1.03230	1.05780

Table 4: Estimated Linear Utility Function Price Indexes for Six Japanese Prefectures

m	$P_{\text{LU}}^{1,m}$	$P_{\text{LU}}^{2,m}$	$P_{\text{LU}}^{3,m}$	$P_{\text{LU}}^{4,m}$	$P_{\text{LU}}^{5,m}$	$P_{\text{LU}}^{6,m}$
1	0.99553	1.00000	1.01042	1.19924	1.14125	1.08775
2	0.98554	1.00496	1.03357	1.16931	1.12093	1.07850
3	0.99018	1.00824	1.01394	1.18753	1.12932	1.08316
4	0.97783	1.00698	1.02535	1.19429	1.13256	1.07406
5	0.99358	1.00680	1.01654	1.20331	1.11740	1.08359
6	0.99335	0.99254	1.01759	1.20800	1.07653	1.07795
7	0.99371	1.01352	1.02757	1.19984	1.13198	1.08235
8	0.98365	1.00819	1.03167	1.19712	1.11301	1.08592
9	0.98762	0.99984	1.00058	1.20745	1.13263	1.09403
10	0.98052	1.00748	1.01232	1.20666	1.13579	1.07562
11	0.98261	0.99631	0.99325	1.18816	1.13190	1.08781
12	0.98437	1.01000	1.01615	1.19104	1.13542	1.08244
13	0.99469	1.00256	0.98728	1.22442	1.13769	1.09945
14	0.98031	1.00379	0.99698	1.21988	1.13007	1.09099
15	0.99059	1.00318	0.99445	1.20359	1.12887	1.10848
16	0.99153	1.01597	1.02436	1.20133	1.12652	1.11148
17	0.98806	1.01585	0.99270	1.20904	1.13251	1.12447
18	0.99382	1.00487	0.99190	1.20293	1.07325	1.09492
19	0.97766	1.00255	1.02362	1.20996	1.13251	1.10249
20	0.99403	1.00233	0.99589	1.20079	1.11509	1.09854
21	1.00723	1.00569	1.02814	1.20168	1.12252	1.08715
22	1.00050	1.01392	1.01723	1.21413	1.13571	1.08135
23	1.00223	1.01523	1.01724	1.23399	1.13922	1.08625
24	1.02286	1.03910	1.04074	1.21932	1.15572	1.10805
Mean	0.99133	1.00750	1.01290	1.20390	1.12620	1.09110

Table 5: Alternative CES Price Indexes for Six Japanese Prefectures

m	$P_{\text{CES}}^{1,m}$	$P_{\text{CES}}^{2,m}$	$P_{\text{CES}}^{3,m}$	$P_{\text{CES}}^{4,m}$	$P_{\text{CES}}^{5,m}$	$P_{\text{CES}}^{6,m}$	$P_{\text{ACES}}^{1,m}$	$P_{\text{ACES}}^{2,m}$	$P_{\text{ACES}}^{3,m}$
1	0.98432	1.00000	1.00859	1.23490	1.13626	1.09615	1.02221	1.00000	0.95253
2	0.97450	1.00654	1.03115	1.20769	1.11630	1.08648	0.96334	1.00433	1.02377
3	0.97981	1.00820	1.01475	1.22291	1.12511	1.09014	1.02102	1.01909	1.03711
4	0.96686	1.00793	1.02686	1.22892	1.12923	1.08164	0.98921	0.99693	1.04497
5	0.97998	1.00797	1.01610	1.23564	1.11337	1.08909	1.01483	1.01220	1.00386
6	0.98136	0.99190	1.01893	1.24099	1.07489	1.08429	1.01529	0.98922	1.03095
7	0.98219	1.01295	1.02692	1.23363	1.12592	1.08903	1.01588	1.00678	1.04099
8	0.97276	1.00801	1.03029	1.23099	1.10804	1.09195	0.99992	1.00666	1.02105
9	0.97689	1.00016	1.00219	1.23861	1.12796	1.10140	1.00887	0.96626	1.02597
10	0.96968	1.00918	1.00981	1.24066	1.12967	1.08205	0.98927	1.01404	0.95669
11	0.97188	0.99717	0.99217	1.22095	1.12424	1.09206	0.90153	0.96618	0.94712
12	0.97328	1.01185	1.01646	1.22594	1.12712	1.08673	1.00142	1.00690	0.97605
13	0.98380	1.00406	0.99031	1.25795	1.12887	1.10299	1.01577	0.96520	0.97867
14	0.96984	1.00438	0.99663	1.25243	1.12158	1.09500	1.00134	0.94197	0.97065
15	0.98062	1.00418	0.99714	1.23825	1.11918	1.11312	1.02316	1.01364	1.02781
16	0.98207	1.01942	1.02555	1.23506	1.11559	1.11504	0.82039	1.03531	1.04858
17	0.97683	1.01780	0.99163	1.24388	1.12152	1.12711	0.95689	1.03614	0.95729
18	0.98308	1.00747	0.99440	1.24032	1.06572	1.09789	1.02362	1.03051	0.88897
19	0.96685	1.00385	1.02560	1.24251	1.12152	1.10658	0.97722	1.02052	1.04532
20	0.98214	1.00258	0.99484	1.23671	1.10441	1.10068	1.01541	1.01277	0.96198
21	0.99568	1.00533	1.02667	1.23947	1.11284	1.09067	1.01046	0.96695	1.04094
22	0.98894	1.01338	1.01603	1.25292	1.12499	1.08401	0.99314	0.96685	0.96149
23	0.99144	1.01532	1.01610	1.26748	1.12850	1.09023	1.01625	0.98478	0.99795
24	1.01085	1.03944	1.04031	1.25336	1.14438	1.11315	1.04092	1.01765	1.03942
Mean	0.98024	1.00830	1.01290	1.23840	1.11860	1.09610	0.99322	0.99920	0.99917

Table 5: Continued

m	$P^{1,m}_{\text{CCES}}$	$P^{2,m}_{\text{CCES}}$	$P^{3,m}_{\text{CCES}}$	$P^{4,m}_{\text{CCES}}$	$P^{5,m}_{\text{CCES}}$	$P^{6,m}_{\text{CCES}}$	$P^{1,m}_{\text{ACCES}}$	$P^{2,m}_{\text{ACCES}}$	$P^{3,m}_{\text{ACCES}}$
1	1.26727	1.18477	1.13152	0.97613	1.00000	1.02977	1.38471	1.25662	1.20390
2	1.24586	1.15494	1.10371	0.96912	1.01454	1.03509	1.34748	1.24787	1.19390
3	1.25819	1.16647	1.12117	0.97004	1.01424	1.03082	1.37113	1.24875	1.19819
4	1.26216	1.17738	1.12287	0.96519	1.01299	1.04189	1.37996	1.25381	1.19733
5	1.26551	1.15414	1.10714	0.94453	1.01553	1.02458	1.38537	1.21824	1.18678
6	1.28222	1.10347	1.11360	0.94480	1.00184	1.02937	1.37905	1.20810	1.19219
7	1.26853	1.16364	1.04533	0.94942	1.01753	1.03620	1.36776	1.21816	1.17012
8	1.26334	1.14991	1.10896	0.94713	1.01297	1.04001	1.36662	1.21280	1.19580
9	1.26526	0.50269	1.11678	0.96147	1.00508	1.02518	1.36185	1.12510	1.20122
10	1.26600	1.07336	1.03876	0.94136	1.01433	1.01722	1.36478	1.21128	1.19435
11	1.23520	1.16356	1.11916	0.94057	0.99860	1.00934	1.35830	1.21554	1.19372
12	1.27576	1.16207	1.12434	0.95567	1.01674	1.01715	1.36982	1.21061	1.19886
13	1.26707	1.16654	1.14257	0.96230	1.00420	1.00432	1.38776	1.21423	1.20744
14	1.27918	1.15918	1.11597	0.94058	0.99968	1.00068	1.38532	1.20875	1.19372
15	1.26798	1.14557	1.13636	0.96108	1.00591	1.00558	1.38239	1.19228	1.20446
16	1.26763	1.13747	1.11802	0.95900	1.02190	1.02546	1.35990	1.19242	1.19296
17	1.27143	1.09282	1.09068	0.96012	1.02125	1.00703	1.37479	1.19697	1.20292
18	1.19816	0.92140	1.08874	0.95966	1.01070	0.99692	1.33685	1.14510	1.18242
19	1.17200	1.13599	1.10831	0.92480	1.00979	1.02111	1.33337	1.19477	1.19106
20	1.26324	1.12905	1.11147	0.96456	1.00138	1.00935	1.36014	1.19158	1.19051
21	1.24511	1.14561	1.07811	0.97529	1.00536	1.02677	1.36031	1.20150	1.17880
22	1.26725	0.85702	1.09922	0.95889	1.01558	1.03898	1.36358	1.16385	1.18541
23	1.28775	1.16122	1.07832	0.96566	1.01766	1.03212	1.38418	1.21098	1.18134
24	1.24154	1.15567	1.10707	0.99066	1.03616	1.05040	1.36587	1.21366	1.19571
Mean	1.25770	1.09850	1.10530	0.95783	1.01140	1.02301	1.36800	1.20640	1.19300

Table 6: Expenditure Ratios for Five Prefectures Relative to Tokyo Prefecture for Feenstra's Method for Adjusting For Variety Changes

m	$\lambda^{1,m}$	$\lambda^{3,m}$	$\lambda^{4,m}$	$\lambda^{5,m}$	$\lambda^{6,m}$	$\mu^{1,m}$	$\mu^{3,m}$	$\mu^{4,m}$	$\mu^{5,m}$	$\mu^{6,m}$
1	1.20231	1.07822	1.00000	1.08543	1.00000	0.79849	0.82257	0.26450	0.31162	0.46641
2	1.18844	1.07353	1.00000	1.09716	1.00000	0.79358	0.82006	0.23832	0.30275	0.47014
3	1.19997	1.06358	1.00000	1.10777	1.00000	0.77661	0.82382	0.26155	0.30002	0.46938
4	1.19303	1.07364	1.00000	1.11060	1.00000	0.78831	0.80790	0.26793	0.27680	0.48433
5	1.23758	1.13580	1.00000	1.08930	1.00000	0.78909	0.81245	0.26517	0.38268	0.47729
6	1.36470	1.13256	1.00000	1.08584	1.00000	0.79424	0.79014	0.26529	0.37184	0.50640
7	1.34626	1.10605	1.00000	1.09260	1.00000	0.79145	0.83769	0.27317	0.39198	0.48750
8	1.37296	1.07043	1.00000	1.09562	1.00000	0.64519	0.85004	0.28657	0.38571	0.49214
9	1.34593	1.06642	1.00000	1.09667	1.00000	0.65514	0.86320	0.28067	0.37657	0.49084
10	1.37275	1.07395	1.00000	1.09947	1.00000	0.66197	0.86026	0.28324	0.41061	0.48937
11	1.36980	1.06994	1.00000	1.07709	1.00000	0.67703	0.86202	0.27959	0.40356	0.49514
12	1.34770	1.08799	1.00000	1.10268	1.00000	0.67166	0.87504	0.28321	0.38786	0.52102
13	1.34604	1.08472	1.00000	1.08845	1.00000	0.66844	0.85971	0.27801	0.39167	0.50904
14	1.37442	1.06866	1.00000	1.10779	1.00000	0.66440	0.86941	0.29237	0.40959	0.50344
15	1.34845	1.06502	1.00000	1.07682	1.00000	0.67109	0.86962	0.28589	0.44689	0.49567
16	1.32884	1.07675	1.00000	1.10389	1.00000	0.66019	0.86987	0.27220	0.41673	0.50382
17	1.30928	1.06781	1.00000	1.08723	1.00000	0.65245	0.85578	0.29067	0.42333	0.50436
18	1.31396	1.06561	1.00000	1.06356	1.00000	0.65713	0.86027	0.29014	0.41691	0.52151
19	1.35678	1.07028	1.00000	1.08673	1.00000	0.65684	0.85865	0.29830	0.42970	0.52021
20	1.31148	1.06358	1.00000	1.07660	1.00000	0.66173	0.85940	0.29923	0.42263	0.50508
21	1.31914	1.06800	1.00000	1.07282	1.00000	0.65360	0.87339	0.29904	0.43350	0.49356
22	1.34536	1.06629	1.00000	1.08263	1.00982	0.64132	0.86317	0.33337	0.47230	0.48144
23	1.34607	1.06195	1.00000	1.08113	1.01199	0.64699	0.83514	0.31800	0.45229	0.46938
24	1.34042	1.05930	1.00000	1.08833	1.00847	0.66521	0.85660	0.31780	0.44712	0.49423

Table 7: Two CES Based Measures of the Increase in the Cost of Living for Five Prefectures due to Differences in Product Availability Relative to the Tokyo Prefecture

m	$D_{\text{CES}}^{1,m}$	$D_{\text{CES}}^{3,m}$	$D_{\text{CES}}^{4,m}$	$D_{\text{CES}}^{5,m}$	$D_{\text{CES}}^{6,m}$	$D_{\text{CCES}}^{1,m}$	$D_{\text{CCES}}^{3,m}$	$D_{\text{CCES}}^{4,m}$	$D_{\text{CCES}}^{5,m}$	$D_{\text{CCES}}^{6,m}$
1	1.00185	1.00545	1.06207	1.05031	1.03514	1.00744	1.02206	1.27357	1.21788	1.14876
2	1.00265	1.00579	1.06709	1.05117	1.03476	1.01070	1.02344	1.29794	1.22190	1.14710
3	1.00320	1.00600	1.06261	1.05114	1.03484	1.01290	1.02432	1.27618	1.22177	1.14744
4	1.00278	1.00646	1.06145	1.05486	1.03337	1.01122	1.02621	1.27059	1.23923	1.14091
5	1.00107	1.00365	1.06195	1.04041	1.03406	1.00432	1.01472	1.27299	1.17247	1.14396
6	0.99636	1.00504	1.06192	1.04192	1.03129	0.98546	1.02040	1.27289	1.17929	1.13171
7	0.99713	1.00346	1.06052	1.03914	1.03307	0.98853	1.01397	1.26613	1.16672	1.13956
8	1.00550	1.00428	1.05822	1.03977	1.03262	1.02229	1.01731	1.25515	1.16956	1.13760
9	1.00571	1.00376	1.05922	1.04086	1.03275	1.02314	1.01517	1.25991	1.17447	1.13815
10	1.00434	1.00359	1.05878	1.03667	1.03289	1.01756	1.01450	1.25782	1.15559	1.13877
11	1.00342	1.00367	1.05940	1.03845	1.03234	1.01380	1.01481	1.26079	1.16358	1.13634
12	1.00452	1.00223	1.05878	1.03921	1.02996	1.01828	1.00898	1.25785	1.16701	1.12586
13	1.00479	1.00317	1.05967	1.03936	1.03105	1.01939	1.01278	1.26209	1.16770	1.13064
14	1.00412	1.00333	1.05726	1.03643	1.03156	1.01665	1.01346	1.25058	1.15453	1.13291
15	1.00453	1.00348	1.05833	1.03367	1.03229	1.01833	1.01405	1.25569	1.14225	1.13612
16	1.00595	1.00297	1.06069	1.03578	1.03153	1.02409	1.01197	1.26695	1.15164	1.13276
17	1.00716	1.00409	1.05754	1.03576	1.03148	1.02906	1.01653	1.25191	1.15154	1.13254
18	1.00667	1.00395	1.05763	1.03751	1.02992	1.02706	1.01594	1.25233	1.15937	1.12567
19	1.00523	1.00383	1.05630	1.03508	1.03003	1.02117	1.01548	1.24603	1.14851	1.12618
20	1.00644	1.00408	1.05615	1.03630	1.03141	1.02611	1.01648	1.24532	1.15395	1.13225
21	1.00674	1.00316	1.05618	1.03527	1.03249	1.02733	1.01273	1.24547	1.14937	1.13700
22	1.00670	1.00376	1.05100	1.03084	1.03319	1.02720	1.01520	1.22109	1.12972	1.14013
23	1.00628	1.00545	1.05324	1.03292	1.03428	1.02545	1.02207	1.23162	1.13893	1.14495
24	1.00521	1.00441	1.05327	1.03315	1.03203	1.02107	1.01783	1.23176	1.13994	1.13498
Mean	1.00410	1.00410	1.05870	1.03940	1.03240	1.01660	1.01670	1.25760	1.16820	1.13680

Table 8: KBF Price Indexes for Six Japanese Prefectures

m	$P_{\text{KBF}}^{1,m}$	$P_{\text{KBF}}^{2,m}$	$P_{\text{KBF}}^{3,m}$	$P_{\text{KBF}}^{4,m}$	$P_{\text{KBF}}^{5,m}$	$P_{\text{KBF}}^{6,m}$
1	1.00750	1.00000	1.01673	1.22156	1.16826	1.16203
2	0.99729	1.00219	1.03100	1.19432	1.14991	1.14966
3	0.99879	1.00764	1.01209	1.21193	1.15530	1.15177
4	0.99029	1.00682	1.02002	1.22187	1.15858	1.13754
5	1.00754	1.00632	1.01320	1.22941	1.14437	1.15480
6	1.00856	0.99199	1.01200	1.23697	1.11237	1.14718
7	1.00882	1.01259	1.02404	1.22846	1.15770	1.14954
8	0.99997	1.00658	1.02667	1.21837	1.14243	1.15528
9	1.00253	0.99920	0.99641	1.23171	1.15935	1.16195
10	0.99762	1.00638	1.02124	1.23365	1.16111	1.14145
11	0.99842	0.99817	1.00641	1.21512	1.15780	1.15846
12	0.99996	1.00885	1.02296	1.21965	1.16225	1.15683
13	1.00881	1.00230	0.99388	1.24911	1.16387	1.17004
14	0.99543	1.00587	1.00436	1.24914	1.15837	1.16224
15	1.00587	1.00278	0.99547	1.23194	1.15813	1.18140
16	1.00493	1.01269	1.02346	1.22637	1.15608	1.18152
17	1.00345	1.01478	1.00598	1.23750	1.15968	1.19261
18	1.00765	1.00284	0.99415	1.22710	1.09765	1.16746
19	0.99100	1.00313	1.02183	1.23460	1.15853	1.17477
20	1.00785	1.00167	1.00710	1.22346	1.14226	1.16712
21	1.02031	1.00602	1.03212	1.22529	1.14791	1.16295
22	1.01379	1.01504	1.03993	1.23543	1.16563	1.14856
23	1.01575	1.01826	1.03262	1.25912	1.16641	1.15835
24	1.03464	1.03978	1.05133	1.24554	1.18268	1.17943
Mean	1.00530	1.00720	1.01690	1.22950	1.15360	1.16140