On Measuring Aggregate Price Changes with Product Turnover

Naohito Abe*, Noriko Inakura†, DS Prasada Rao‡, Akiyuki Tonogi§

Preliminary draft. Please do not cite or circulate.

December 20, 2022

Abstract

Aggregation of prices or quantities when goods disappear or appear has been considered particularly challenging. Product turnover occurs at various aggregation levels. Frequent product turnover is observed in commodity-level scanner data. Various fresh foods, such as fruit, exhibit marked seasonal tendencies. Sometimes, we observe the disappearance of commodities at category levels, as in the case of foreign trips during the COVID-19 pandemic. When we compare price levels across countries, we often encounter differences in product sets. In some countries, beef or pork is not consumed, while in other countries, such meats are quite popular. When commodity sets differ between two time periods, or across countries or regions, the current standard method used to construct the cost-of-living index, a price index number based on economic theory, is the method proposed by Feenstra (1994). While the Feenstra index that accounts for various effects has been widely used, two problems remain. First, the Feenstra index is subject to chain drifts, which makes it difficult to compare price levels between two remote periods. The second problem is that the index cannot be defined when the elasticity of substitution is less than or equal to unity, which makes the index inapplicable for aggregation at the category level. In this paper, we propose a cost-of-living index when product turnover occurs, which is free from chain drift, and that can be applied when the elasticity of substitution is small. Our proposed cost-of-living index can also be applied when preferences vary over time or across regions. The recently proposed Redding and Weinstein's (2020) unified cost-of-living index can handle product turnover and preference heterogeneity, but their index is subject to chain drift because product turnover is assumed to occur purely due to supply shocks. In contrast to Redding and Weinstein's index, which assumes cardinal utility, our index is based on ordinal utility. The assumption of ordinal utility enables us to consider product turnover caused not only by supply shocks but also by demand shocks. We will show that when the product turnover is caused by demand shocks, the cost-of-living index is free from chain drift. Our index can also be applied to price data at different levels of aggregation: ranging from commodity-level scanner data to international price comparisons above the elementary level. We also provide a procedure to decompose

*Hitotsubashi University, nabe@ier.hit-u.ac.jp
†Shikoku University
‡University of Queensland
§Toyo University

price changes due to demand and supply shocks. Our index exhibits notable differences from various indexes when applied to weekly scanner data. For example, in the case of ice cream for which there are strong seasonal trends in purchasing patterns and varieties, our index goes up during the summer, which makes the "real expenditures" on ice cream much smoother than the nominal expenditures, or other real expenditures obtained using the Fisher's index, Feenstra's index, and the CUPI (CES-Unified Index) by Redding and Weinstein (2020) as deflators. Furthermore, the CUPI and Feenstra's index show a strongly negative drift, while our index is transitive and exhibits little drift. Other applications of our index to national-level data as well as international price comparisons data are also provided.

1 Introduction

Thanks to widespread usage of transaction data via scanner and many "new" data sources through internet, aggregation of quantities or prices has become an increasingly important topic in economic measurement and analysis. The usual approach to obtain an aggregate quantity is through the real expeuditure or consumption. More specifically, when aggregating several quantities, we deflate the total nominal expenditures over the basket of goods and services using the consumer price index (CPI). The resulting real consumption can be regarded as a single commodity. The index number theory implies that as long as the CPI is equivalent with the cost of living index (COLI), the real consumption equals the utility level, which makes real consumption an appropriate aggregate. When the product set remains constant over time or regions, that is, if the variety of commodities are identical over the time or regions, superlative indexes such as Fisher, Tornqvist, and Walsh are known to provide a good approximation for the COLI. However, when product set changes over time or regions, the superlative index is not an appropriate measure of the cost of living. In other words, the real consumption obtained using a superlative index as deflator is no longer the utility level.

Differences in commodity sets appear at different levels of aggregation, from high frequency scanner data on transaction prices to international price comparisons at category level data. Scanner data of daily or weekly transactions often show frequent product turnover. Figure 1 shows the number of new ice creams, as well as their sales share, introduced to Japanese market after the first week of January, 2017. Figure 2 illustrates the seasonal pattern of the amount of sales and the degree of product variety. As is clear from the figures, during summer, the product variety tends to increase while in winter, the number of products tends to decrease. According to Figure 1, within 6 years, only 30-40 % of the commodities survive. In other products such as potato chips, the ratio becomes greater than 90 % . This implies that the direct price index that compares prices between 2022 and 2017, reflects just a part of the actual movements in prices. Even at a country level, during the lock down period caused by COVID19-Pandemic, we observe disappearance of some product categories, such as international tour packages, and admission tickets to professional sport games, in monthly data. Another example of the category level difference is beef and pork in international comparisons. The consumption data of beef and pork do not exist in many Middle and Near East, and South Asian countries while such items are popular in other Asian countries. Another example of product

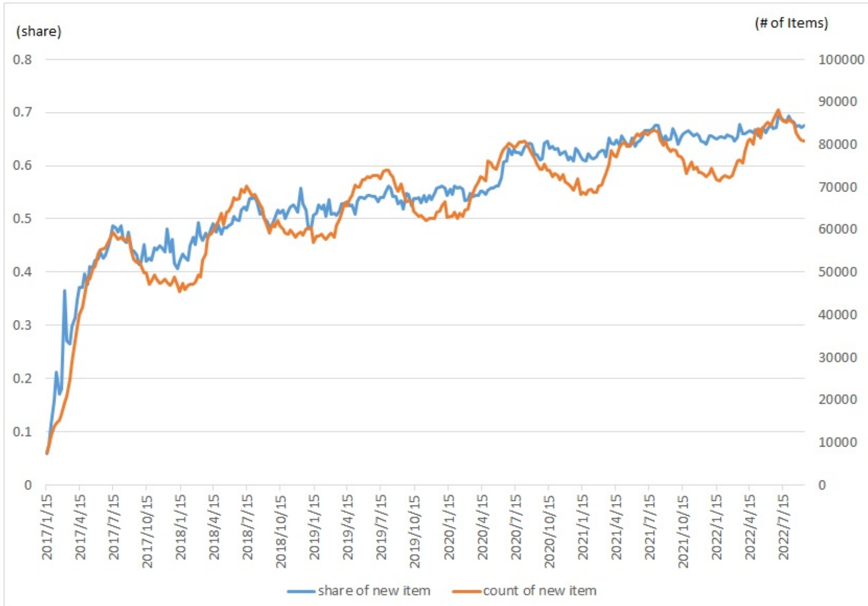

Figure 1: Sales Share and Number of New Items (Ice Cream)

turnover comes from seasonal products such as fresh fruits.

If preferences do not vary over time or if preferences are identical across regions, the standard method to deal with different commodity sets is to impute reservation prices proposed by Diewert and Fox (2022). Diewert and Fox (2022) argued that when imputing prices for disappearing products, we should use reservation prices at which demands for missing products become zero. A problem in using reservation price is that if the marginal utility does not reach zero at any finite level of prices, the reservation prices goes infinity. However, in a path-breaking paper, Feenstra (1994) has shown that as long as the utility function is the constant elasticity of substitution (CES), and the elasticity of substitution is greater than unity, the aggregate price index with infinite prices becomes the Sato-Vartia index with the variety effects. Because the variety effect index by Feenstra (1994) is easy to understand and build, the Feenstra index has become the standard COLI when product turnover occurs. [1] However, Feenstra (1994) index has several limitations. First, it is subject to chain drift, which makes it difficult to compare price levels between two remote periods. Figure 3 shows the chain drift associated with the Jevons and Feenstra indices applied to Japanese ice cream data.[2] Note that the Jevons index is transitive if the product variety is constant over time. This implies that the discrepancies between the direct and chained Jevons indexes are caused by the changes in the product variety of ice cream. The second problem is that the Feenstra (1994) index cannot be defined when the elasticity of substitution is less than or equal to unity. These issues makes the Feenstra index inapplicable for aggregation at the category level. That is, we cannot analyze the effects of the disappearance of product categories using the Feenstra index. The Feestra's variety effect index is the only cost of living index in the literature that can account for changes in the commodity sets. In conclusion, currently, there is no good

[1] See Hottman et al. (2016) and Redding and Weinstein (2020) for recent applications.

[2] The detail of the figure and data is discussed in Section7

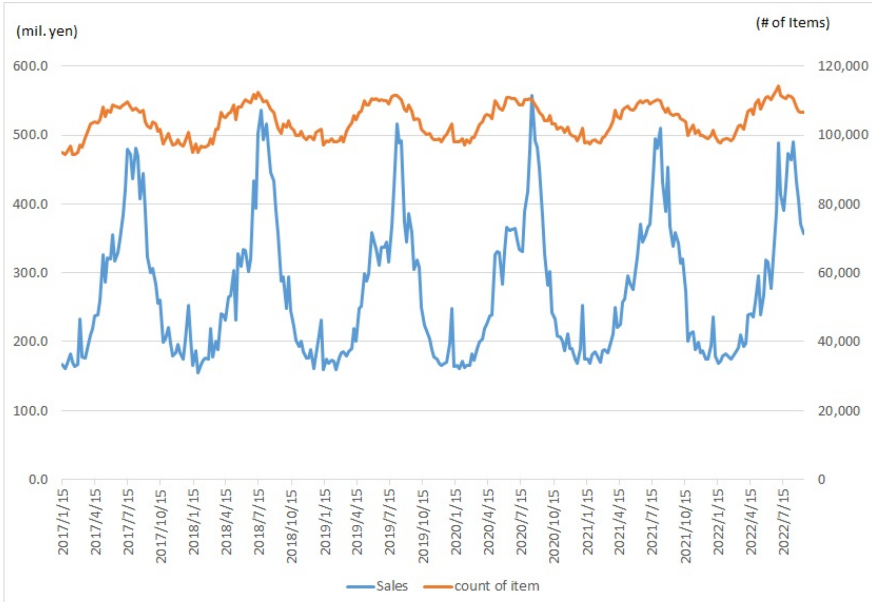

Figure 2: Sales and the Number of Different Commodities (Ice Cream)

method to evaluate changes in cost of living in the presence of changing commodity sets .

In this paper, we propose a cost-of-living index that can be easily applied when product turnover occurs, which is free from chain drift, and that can be applied to category level aggregation whose elasticity of substitution is less than unity. Our proposed cost-of-living index can also be applied when preferences vary over time or across regions. The main idea behind our proposed method is as follows. When people's taste/preference for a commodity becomes zero, the demand for the product also becomes zero, which leads to disappearance of the product. This mechanism is suitable when conducting international comparisons where preferences vary across countries. This index can also be applied to seasonal products such as ice cream whose demand in winter is smaller than during summer. A complexity we face when making a cost of living index with heterogeneous preferences is that we need to compare expenditure functions under different utility functions. In other words, we need to conduct interpersonal comparisons of ordinal utilities. While we can compare whether a person is happier or less happier than other at a given quantity vector, the result critically depends on the selection of the reference quantity vector. In this paper, we show that just two popular axioms that are widely used in index number theory enables us to characterize the reference quantity vector uniquely.

How does our proposed index compare with the Redding and Weinstein (2020) unified cost-of-living index? While their index can handle product turnover and preference heterogeneity, their index is subject to chain drift. Further, the Redding and Weinstein (2020)'s index assumes cardinal utility. In contrast our index is based on ordinal utility. The assumption of ordinal utility enables us to consider product turnover caused not only by supply shocks but also by demand shocks. We show that when product turnover is caused by demand shocks, the cost-of-living index is free from chain drift. We will also show that our index can be

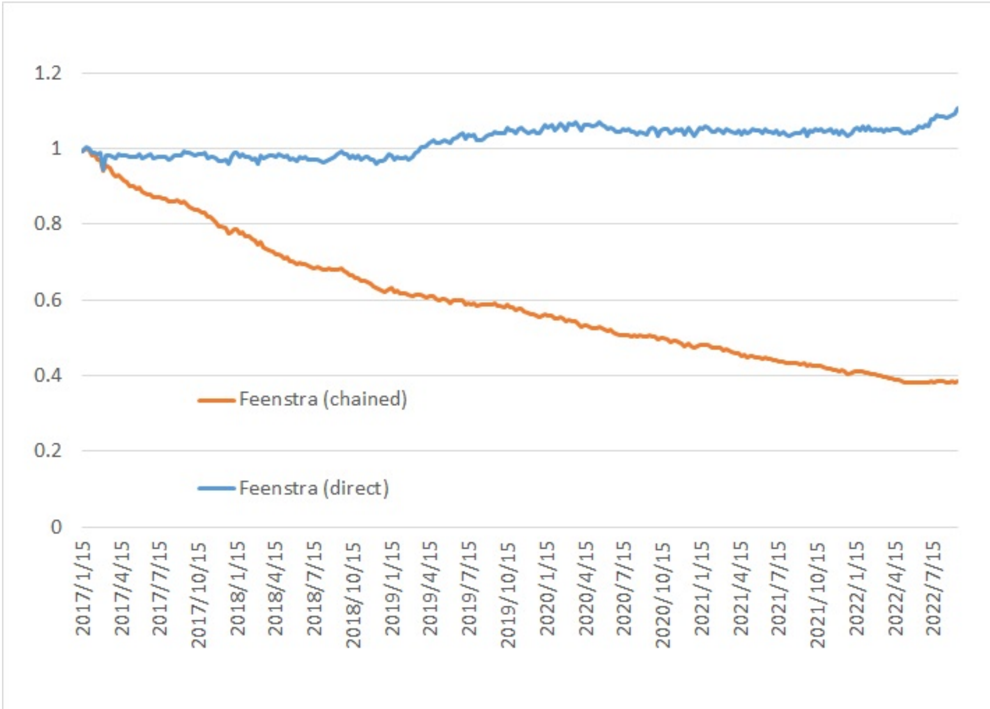

Figure 3: Chain Drift of Feenstra's Index

applied to price data at different levels of aggregation: from high frequency commodity-level scanner data to international comparisons. Another contribution of the paper is that it provides a method to decompose price changes due to demand and supply shocks. Our proposed index exhibits notable differences from various indexes when applied to weekly scanner data. For example, in the case of ice cream for which there are strong seasonal trends in purchasing patterns and varieties, our index goes up during the summer, which makes the "real expenditures" on ice cream that are smoother than the observed nominal expenditures, or real expenditures obtained using the Fisher index, Feenstra index, and the CUPI (CES-Unified Index) by Redding and Weinstein (2020) as delfators. Furthermore, the CUPI and Feenstra index show a strongly negative drift, whereas our index is transitive with little drift. The paper presents empirical results based on the application of our approach to national-level data and cross-country data used for international price and real expenditure comparisons.

2 Product Turnover and The Impossibility of Making a Transitive Bilateral Price Index

In this section, we briefly demonstrate the difficulty of constructing a price index when product turnover occurs. The classical work by Funke et al. (1977) shows that any bilateral index with basic properties such as monotonicity and identity, cannot be transitive unless the index is a Cobb-Douglas function of price relatives. When the product set changes over time, even if we relax the requirements for the price index such that it does not need to pass the identity or monotonicity tests, we cannot construct a bilateral price index that is transitive. The impossibility is easy to show as follows.

Consider three time periods, s, t, and k and the case where the commodity sets vary over time. Let Ω_{st} denote the set of commodities that exist both at time s and t, while Ω_s is the set of all the commodities that exist at time s. Thus, we have

$$\Omega_s \cap \Omega_t = \Omega_{st}$$

We assume

$$\Omega_s \cap \Omega_t \cap \Omega_k \neq \phi,$$

that is, at least one commodity exists in all three periods.

We also assume that a commodity, say i, exists in Ω_{st}, but not in Ω_{tk} nor Ω_{sk}. In other words, commodity i exists only at times s and t, but not at k. Let p_k^{st} be the price vector of goods that exist in both time t and s at time k.

Suppose the price index between time s and $t, P(s,t)$, depends only on prices and quantities of common goods, Ω_{st}, only. Thus, the price index can be written as:

$$P(s,t) = P\left(p_s^{st}, p_t^{st}, q_s^{st}, q_t^{st}\right) \tag{1}$$

We also assume that if the price of a commodity in Ω_{st} changes, the price index also changes. More precisely, we have

$$P\left(p_s^{st}, p_t^{st}, q_s^{st}, q_t^{st}\right) \neq P\left(p_s^{st}, p_t^{*st}, q_s^{st}, q_t^{st}\right) \tag{2}$$

where

$$p_t^{st} = \left(p_{1,t}^{st}, p_{2,t}^{st}, p_{3,t}^{st}, .., p_{i,t}^{st}, .., p_{Nst,t}^{st}\right) \tag{3}$$

$$p_t^{*st} = \left(p_{1,t}^{st}, p_{2,t}^{st}, , .p_{i,t}^{*st}, .., p_{Nst,t}^{st}\right) \tag{4}$$

$$p_{i,t}^{st} \neq p_{i,t}^{*st} \tag{5}$$

Proposition 1 *$P(s,t)$ in equation (1) is not transitive. That is, for any $P(s,t)$, we can find price and quantity vectors such that the following inequality holds,*

$$P\left(p_s^{st}, p_t^{st}, q_s^{st}, q_t^{st}\right) \times P\left(p_t^{tk}, p_k^{tk}, q_t^{tk}, q_k^{tk}\right) \neq P\left(p_s^{sk}, p_k^{sk}, q_s^{sk}, q_k^{sk}\right). \tag{6}$$

Proof.

Suppose the following equality holds for all quantity and price vectors,

$$P\left(p_s^{st}, p_t^{st}, q_s^{st}, q_t^{st}\right) \times P\left(p_t^{tk}, p_k^{tk}, q_t^{tk}, q_k^{tk}\right) = P\left(p_s^{sk}, p_k^{sk}, q_s^{sk}, q_k^{sk}\right). \tag{7}$$

then , change the price of commodity i that exists only at time s and t. By assumption, $P\left(p_s^{st}, p_t^{st}, q_s^{st}, q_t^{st}\right)$ will have a different value while the other price index numbers, $P\left(p_t^{tk}, p_k^{tk}, q_t^{tk}, q_k^{tk}\right)$ and $P\left(p_s^{sk}, p_k^{sk}, q_s^{sk}, q_k^{sk}\right)$ remain unchanged, which contradicts (7) and hence transitivity. ■

The above proposition implies that when commodity sets vary over time, a standard bilateral price or quantity index cannot be transitive; and therefore, is subject to chain drift. To obtain a transitive index free from chain drift, the domain of the price index between s and t, should not be Ω_{st}. A practical method to obtain a transitive index is to use the GEKS (Gini- Eltető - Köves -Szulc).[3] If there are M periods, the GEKS is defined as follows,

$$P_{st}^{GEKS} = \prod_{l=1}^{M} \left(P_{sl}^F \times P_{lt}^F\right)^{\frac{1}{M}} \tag{8}$$

where P_{lt}^F is the Fisher index between time l and t while M is the number of time periods in the data set.[4] Although GEKS has been widely used in the International Comparison Program (ICP) at the World Bank[5] and national statistical offices in some countries such as Australia and Italy, it is also known that GEKS is not generally the cost of living index.[6] Although the Fisher index is a superlaitve index, its GEKS is not a superlative index. The lack of an economic model behind GEKS implies that the real expenditure obtained by it does not have an economic interpretation.

3 Cost of Living Index with Product Turnover: The Reservation Price Approach

The cost of living index (COLI) is defined as the ratio of two expenditure functions at two periods given same preferences in both periods. When product sets change over time, the definition of "the same preferences" becomes ambiguous. Suppose the product set at time t is given by Ω_t. Feenstra (1994) and Balk (1999) propose the following COLI

$$COLI\_FB = \frac{E\left(p_1, U\left(\Omega_1\right)\right)}{E\left(p_0, U\left(\Omega_0\right)\right)} \tag{9}$$

where $U\left(\Omega_t\right)$ is the utility level given the commodity set at time t. Recently, Diewert and Fox (2022) introduced the following COLI,

[3] See Diewert (2013) for a description of the method and its properties.
[4] Another method to obtain a transitive index is to use Dutot's index whose domain is the union of Ω_s and Ω_t.
[5] See Rao (2013) for a description of the framework for ICP.
[6] See Neary (2004) for detail.

$$COLI\_DF = \frac{E\left(\widetilde{p}_1, U\left(\Omega\right)\right)}{E\left(\widetilde{p}_0, U\left(\Omega\right)\right)} \tag{10}$$

where $\Omega = \Omega_0 \cup \Omega_1$, $\widetilde{p}_t$ is the price vector for the commodity in the set Ω. When some products disappear at time 1 so that we have $\Omega_1 \neq \Omega_0$, some prices are not observable. Diewert and Fox (2022) advocate the use of reservation prices at which the demand for the commodity that does not exist at time 1 becomes zero. For example, when a commodity disappears at time 0 , in (10) , the price of the disappeared good becomes very high, which increases the COLI_DF. Under the same circumstances, the COLI by (9) also increases due to smaller variety. These two COLIs are generally different. For example, suppose that $U\left(\Omega\right)$ is a Cobb-Douglas function. Also suppose that at time 1, one of the commodities disappears. Then, (10) becomes zero regardless of the other prices at time 1 while (9) takes a finite value. Therefore, the utility functions generally gives us different cost of living index numbers. However, we can show that when we have the CES utility function with the elasticity of substitution that is greater than unity, the two COLIs are always identical.

3.1 Disappearance of Products

In this subsection, we consider the implications of the disappearance of commodities for the cost of living index. We assume that the utility function is of the class of constant elasticity of substitution (CES) as follows;

$$U_t = \left(\sum_{i=1}^{N} \left(\varphi_i q_{it}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \tag{11}$$

where $\sigma > 0$ is the elasticity of substitution. q_{it} is the quantity of commodity i at time $t, \varphi_i \geq 0$ is a parameter that affects the marginal utility of commodity i.

The cost of living index for the CES preference is given by

$$COLI\left(s,t\right) = \frac{\left(\sum_{i=1}^{N}\left(\frac{p_{it}}{\varphi_i}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}}{\left(\sum_{i=1}^{N}\left(\frac{p_{is}}{\varphi_i}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}} \tag{12}$$

From the demand function, we can derive the following relationships

$$w_{it} = \left(\frac{p_{it}}{\varphi_i P_t}\right)^{1-\sigma} \tag{13}$$

$$P_t = \left(\sum_{i=1}^{N}\left(\frac{p_{it}}{\varphi_i}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \tag{14}$$

where w_{it} is the expenditure share of commodity i at time t. p_{it} is the commodity price; and P_t is the unit expenditure function.

Suppose there are two goods, $N = 2$, and two time periods, 0 and 1. In addition, suppose that commodity 2 becomes unavailable at time 1 while the price of the first commodity is unchanged. In the reservation price approach, this disappearance is equivalent for the consumer with the case in which the price of the second commodity becomes extremely high. Because the demand function for commodity i does not have a finite upper limit, we assume that the reservation price of the second commodity is infinite: $p_{21} = \infty$, while $p_{10} = p_{11} < \infty$. As long as the elasticity of substitution, σ, is strictly greater than unity, we obtain

$$\begin{aligned} P_1^{1-\sigma} &= \lim_{p_{21}\to\infty} \left(\left(\frac{p_{11}}{\varphi_1}\right)^{1-\sigma} + \left(\frac{p_{21}}{\varphi_2}\right)^{1-\sigma} \right) \\ &= \left(\frac{p_{11}}{\varphi_1}\right)^{1-\sigma}. \end{aligned} \tag{15}$$

The price of the first commodity is the same in both periods. Thus, the price index for the continuing goods is unity. The cost of the living index becomes

$$\begin{aligned} \frac{P_1}{P_0} &= \frac{\frac{p_{10}}{\varphi_1}}{\left(\left(\frac{p_{10}}{\varphi_1}\right)^{1-\sigma} + \left(\frac{p_{20}}{\varphi_2}\right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}} \\ &= \frac{1}{\left(1 + \left(\frac{\varphi_1}{p_{10}}\frac{p_{20}}{\varphi_2}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}} \end{aligned} \tag{16}$$

Using the following relationships,

$$w_{i0} = \left(\frac{p_{i0}}{\varphi_i P_0}\right)^{1-\sigma} \text{ for } i = 1, 2, \tag{17}$$

we can obtain,

$$\left(\frac{\varphi_1}{p_{10}}\frac{p_{20}}{\varphi_2}\right)^{1-\sigma} = \left(\frac{p_{20}}{\varphi_2 P_0}\right)^{1-\sigma} \times \left(\frac{\varphi_1 P_0}{p_{10}}\right)^{1-\sigma} = \frac{w_{20}}{w_{10}} \tag{18}$$

Therefore, the cost of living index becomes a simple function of the expenditure share of continuing goods at time 0;

$$\begin{aligned} \frac{P_1}{P_0} &= \frac{1}{\left(1 + \left(\frac{\varphi_1}{p_{10}}\frac{p_{20}}{\varphi_2}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}} \\ &= \frac{1}{\left(1 + \frac{w_{20}}{w_{10}}\right)^{\frac{1}{1-\sigma}}} \\ &= w_{10}^{\frac{1}{1-\sigma}} \end{aligned} \tag{19}$$

It is easy to show that the last term is equivalent to the variety effects term in Feenstra (1994). The lambda ratio in Feenstra (1994) is;

$$\frac{\lambda_1}{\lambda_0} = \frac{1}{w_{10}} \tag{20}$$

Thus, the variety effects term in Feenstra (1994) is given by,

$$\left(\frac{\lambda_1}{\lambda_0}\right)^{1/(\sigma-1)} = w_{10}^{\frac{1}{1-\sigma}}, \tag{21}$$

which is identical to that of COLI (19). Note that the above calculation critically relies on the assumption that the elasticity of substitution, σ, is strictly greater than unity. If the elasticity is equal to unity, the COLI becomes infinite when the product disappears.

3.2 Product Entry

Product entry is the reverse of the previous case of product disappearance. Suppose that a new product appears at time 1, whereas at time 0, there is only one commodity. This case is equivalent to the case in which the price of the new commodity is infinite before entrance. Thus, if the elasticity of substitution is greater than unity, the unit expenditure function at time 0 is given by

$$\begin{aligned} P_0 &= \lim_{p_{20}\to\infty}\left(\left(\frac{p_{10}}{\varphi_1}\right)^{1-\sigma} + \left(\frac{p_{20}}{\varphi_2}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \\ &= \left(\frac{p_{10}}{\varphi_1}\right) \end{aligned} \tag{22}$$

The COLI can be written as

$$\frac{P_1}{P_0} = \left(\left(\frac{p_{11}}{\varphi_1}\right)^{1-\sigma} + \left(\frac{p_{21}}{\varphi_2}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}} / \frac{p_{10}}{\varphi_1} \tag{23}$$

Using the demand function, and assuming $p_{10} = p_{11}$, we can show that

$$\begin{aligned} \frac{P_1}{P_0} &= \left(1 + \left(\frac{p_{21}\varphi_1}{p_{11}\varphi_2}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \\ &= w_{11}^{\frac{1}{\sigma-1}} \end{aligned} \tag{24}$$

Note that the lambda ratio by Feenstra (1994) is given by,

$$\left(\frac{\lambda_1}{\lambda_0}\right)^{1/(\sigma-1)} = w_{11}^{\frac{1}{\sigma-1}} \tag{25}$$

Therefore, the COLI becomes that proposed in Feenstra (1994),

$$\frac{P_1}{P_0} = \left(\frac{\lambda_1}{\lambda_0}\right)^{1/(\sigma-1)} \tag{26}$$

3.3 General Case

Suppose the commodity sets vary between time 0 and 1. We denote the various sets in time 0 and 1 as follows,

$$\Omega = \Omega_0 \cup \Omega_1, \Omega_{01} = \Omega_0 \cap \Omega_1, \Omega_{01}^{1c} = \Omega_1 \backslash \Omega_{01}, \Omega_{01}^{0c} = \Omega_0 \backslash \Omega_{01}$$

The utility function is defined over the entire set, Ω. At time each time, only the subset of the commodities are available. The cost of living index is defined as the ratio of the two expenditure functions, $E(p, U)$, as

$$COLI = \frac{E(p_1, U(\Omega))}{E(p_0, U(\Omega))}, \tag{27}$$

where $\widetilde{p}_t \in R_{++}^{\#(\Omega)}$is the price vector at time t that includes unobservable prices at time t. $U(\Omega)$ is the utility level at which the expenditure function is evaluated.

Under the CES preference structure, the expenditure function at time 1 can be written as

$$E(p_1, U(\Omega)) = U(\Omega)\left(\sum_{i\in\Omega}\left(\frac{\widetilde{p}_{i1}}{\varphi_i}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}. \tag{28}$$

Now, suppose that the unobservable prices at time 1 are set at the reservation level. Under the CES preference structure, the reservation price is positive infinite. This, if σ is greater than unity, we can rewrite (28) as

$$E(p_1, U(\Omega)) = U(\Omega)\sum_{i\in\Omega_1}\left(\frac{\widetilde{p}_{i1}}{\varphi_i}\right)^{1-\sigma}.$$

Thus, the COLI becomes

$$\begin{aligned} COLI &= \frac{E(\widetilde{p}_1, U(\Omega))}{E(\widetilde{p}_0, U(\Omega))} \\ &= \frac{E(\widetilde{p}_1, 1)\, U(\Omega)}{E(\widetilde{p}_0, 1)\, U(\Omega)} \\ &= \left(\frac{\sum_{i\in\Omega_1}\left(\frac{p_{i1}}{\varphi_i}\right)^{1-\sigma}}{\sum_{i\in\Omega_0}\left(\frac{p_{i0}}{\varphi_i}\right)^{1-\sigma}}\right)^{\frac{1}{1-\sigma}} \end{aligned} \tag{29}$$

where p_t is the price of commodities that exist at time t. Note that from the first order condition, for all $i \in \Omega_t$, $t = 0, 1$, we get

$$\left(\frac{p_{it}}{\varphi_i}\right)^{1-\sigma} = P_t^{\sigma-1} w_{it} \tag{30}$$

$$\text{where } E\left(\widetilde{p}_t, 1\right) = P_t \tag{31}$$

Then, as Balk (1999) shows, we can multiply both numerator and denominator of (29) by a part of the numerator, [7]

$$\left(\frac{\sum_{i\in\Omega_1}\left(\frac{p_{i1}}{\varphi_i}\right)^{1-\sigma}}{\sum_{i\in\Omega_0}\left(\frac{p_{i0}}{\varphi_i}\right)^{1-\sigma}}\right)^{\frac{1}{1-\sigma}} = \left(\frac{\sum_{i\in\Omega_1}\left(\frac{p_{i1}}{\varphi_i}\right)^{1-\sigma}\sum_{i\in\Omega_{01}}\left(\frac{p_{i1}}{\varphi_i}\right)^{1-\sigma}}{\sum_{i\in\Omega_{01}}\left(\frac{p_{i1}}{\varphi_i}\right)^{1-\sigma}\sum_{i\in\Omega_0}\left(\frac{p_{i0}}{\varphi_i}\right)^{1-\sigma}}\right)^{\frac{1}{1-\sigma}} \tag{32}$$

Then, using (30) , we can rewrite the above equation as

$$\begin{aligned}
&\left(\frac{\sum_{i\in\Omega_1}\left(\frac{p_{i1}}{\varphi_i}\right)^{1-\sigma}\sum_{i\in\Omega_{01}}\left(\frac{p_{i1}}{\varphi_i}\right)^{1-\sigma}}{\sum_{i\in\Omega_{01}}\left(\frac{p_{i1}}{\varphi_i}\right)^{1-\sigma}\sum_{i\in\Omega_0}\left(\frac{p_{i0}}{\varphi_i}\right)^{1-\sigma}}\right)^{\frac{1}{1-\sigma}} \\
&= \left(\frac{\sum_{i\in\Omega_1} P_1^{\sigma-1} w_{i1}\sum_{i\in\Omega_{01}} P_0^{\sigma-1} w_{i0}\left(\frac{p_{it}}{p_{i0}}\right)^{1-\sigma}}{\sum_{i\in\Omega_{01}} P_1^{\sigma-1} w_{i1}\sum_{i\in\Omega_0} P_0^{\sigma-1} w_{i0}}\right)^{\frac{1}{1-\sigma}} \\
&= \left(\frac{\sum_{i\in\Omega_{01}} w_{i0}\left(\frac{p_{it}}{p_{i0}}\right)^{1-\sigma}}{\sum_{i\in\Omega_{01}} w_{i1}}\right)^{\frac{1}{1-\sigma}}
\end{aligned}$$

Then, define the expenditure share in the common commodity set, Ω_{01}, as follows;

$$w_{i1}^c = \frac{p_{i1}q_{i1}}{\sum_{i\in\Omega_{01}} p_{i1}q_{i1}}, w_{i0}^c = \frac{p_{i0}q_{i0}}{\sum_{i\in\Omega_{01}} p_{i0}q_{i0}}$$

Also, define λ_t as follows,

$$\lambda_1 = \frac{\sum_{i\in\Omega_{01}} p_{i1}q_{i1}}{\sum_{i\in\Omega_1} p_{i1}q_{i1}}, \lambda_0 = \frac{\sum_{i\in\Omega_{01}} p_{i0}q_{i0}}{\sum_{i\in\Omega_1} p_{i0}q_{i0}}$$

Then, it is easy to show that

[7]Balk (1999) does not assume $\sigma > 1$ because he defines (29) as the COLI.

$$
\begin{aligned}
COLI &= \left(\frac{\sum_{i\in\Omega_{01}} w_{i0} \left(\frac{p_{it}}{p_{i0}}\right)^{1-\sigma}}{\sum_{i\in\Omega_{01}} w_{i1}} \right)^{\frac{1}{1-\sigma}} \\
&= \left(\frac{\sum_{i\in\Omega_{01}} \lambda_0 w^c_{i0} \left(\frac{p_{it}}{p_{i0}}\right)^{1-\sigma}}{\lambda_1 \sum_{i\in\Omega_{01}} w^c_{i1}} \right)^{\frac{1}{1-\sigma}} \\
&= \left(\frac{\lambda_1}{\lambda_0}\right)^{\frac{1}{\sigma-1}} \left(\sum_{i\in\Omega_{01}} w^c_{i0} \left(\frac{p_{it}}{p_{i0}}\right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}
\end{aligned}
\tag{33}
$$

This is the product of the Lambda ratio by Feenstra (1994) and the Lloyd-Moulton index that is exact for the CES preference. If we replace the Lloyd-Moulton index by the Sato-Vartia index that is also exact for the CES preference, we can obtain the Feenstra index.

4 Product Turnover Caused by Demand Shocks

Movements of prices and quantities in real economic data often reflect not only supply shocks but also demand shocks. When preferences are homothetic, demand shocks can be captured by changes in the taste parameters, φ.

When comparing two price vectors using expenditure functions under heterogeneous preferences, we need to compare the levels of the two utility functions. There are two options for computing the cost of a living index under preference heterogeneity. The first is cardinal COLI by Redding and Weinstein (2020), and the second is ordinal COLI developed first by Balk (1989).

4.1 Cardinal Utility Approach: Redding and Weinstein

In an influential study, Redding and Weinstein (2020) propose the following cost of living index when preferences vary over time:

$$
COLI^C(s,t) = \frac{E_t(p_t, U_t = U)}{E_s(p_s, U_s = U)}. \tag{34}
$$

The utility functions, U_t and U_s have different parameters. Therefore, by setting $U_t = U = U_s$, we assume that utility is cardinal. When the preferences are CES, the COLI by Redding and Weinstein (2020) (hereafter RW) is defined as:

$$COLI^C(s,t) = \left(\frac{\sum_{i=1}^{N_t} \left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma} [U = U_t]}{\sum_{i=1}^{N_t} \left(\frac{p_{is}}{\varphi_{is}}\right)^{1-\sigma} [U = U_s]} \right)^{\frac{1}{1-\sigma}}$$

$$= \left(\frac{\sum_{i=1}^{N_t} \left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}}{\sum_{i=1}^{N_t} \left(\frac{p_{is}}{\varphi_{is}}\right)^{1-\sigma}} \right)^{\frac{1}{1-\sigma}} \tag{35}$$

where N_k $(k = s, t)$is the number of commodities in Ω_k. The $COLI^C(s,t)$ is not homogeneous of degree zero with respect to the taste parameters,φ_{it} and φ_{is}. In order to identify the COLI, or to determine the index number uniquely, we need to impose an exogenous normalization condition for the taste parameters. RW assume the following normalization condition for common goods, N^{Cst}

$$\frac{1}{N^{C_{st}}} \sum_{i \in \Omega_{C_{st}}} \ln \varphi_{is} = \ln \varphi \text{ for all } s. \tag{36}$$

From the first order condition, we obtain

$$\ln \varphi_{it} = \ln p_{it} - \ln P_t - \frac{1}{(1-\sigma)} \ln w_{it}$$

Using the normalization condition (36),

$$\frac{1}{N^{C_{st}}} \sum_{i \in \Omega_{C_{st}}} \ln \varphi_{it} = -\ln P_t + \frac{1}{N^{C_{st}}} \sum_{i \in \Omega_{C_{st}}} \left(\ln p_{it} - \frac{1}{(1-\sigma)} \ln w_{it} \right) = \ln \varphi \tag{37}$$

Therefore, the COLI can be written as

$$\begin{aligned} &\ln P_t - \ln P_s \\ &= \frac{1}{N^{C_{st}}} \sum_{i \in \Omega_{C_{st}}} \left(\ln p_{it} - \frac{1}{(1-\sigma)} \ln w_{it} \right) - \frac{1}{N^{C_{st}}} \sum_{i \in \Omega_{C_{st}}} \left(\ln p_{is} - \frac{1}{(1-\sigma)} \ln w_{is} \right). \end{aligned} \tag{38}$$

Using

$$w_{it}^{C_{st}} = w_{ic} \times \lambda_s^{-1}$$

we obtain the following RW's index,

we obtain the following RW's index,

$$
\begin{aligned}
&\ln P_t - \ln P_s \\
&= \frac{1}{N^{C_{st}}} \sum_{i \in \Omega_{C_{st}}} \left((\ln p_{it} - \ln p_{is}) - \frac{1}{(1-\sigma)} (\ln w_{it} - \ln w_{is}) \right) \\
&= \frac{1}{N^{C_{st}}} \sum_{i \in \Omega_{C_{st}}} \left((\ln p_{it} - \ln p_{is}) - \frac{1}{(1-\sigma)} \left(\ln w_{it}^{C_{st}} - \ln w_{is}^{C_{st}} \right) \right) - \frac{1}{(1-\sigma)} (\ln \lambda_t - \ln \lambda_s) \\
&= \ln CCV + \text{Lambda Effects}
\end{aligned}
$$

where CCV (CES Common Variety) is the price index by RW for common goods. There are several notable characteristics of RW. While RW introduce demand shocks for common goods, product entry and exit are assumed to be solely caused by supply shocks. Second, we cannot compute the RW index when the elasticity of substitution is unity, even if there is no product entry or exit. Third, by taking the geometric mean of the common goods in the normalization condition, (36), the information of the prices of not common goods are ignored. However high or low the prices are, such information are discarded in RW's index. Finally, the index depends on the selection of the normalization condition, (36). RW take the average over the common goods. Alternatively, we can take the average over the smaller subset of common goods, which gives us a different index number but one that is the COLI for the CES preferences.

What happens if the taste parameter φ_{it} becomes zero at time t? Specifically, we assume that people are prevented from buying commodity i at any prices. To simplify the situation, suppose product i is the only disappearing good. No new goods exist at time t. The demand for commodity i at time s is positive, so that $w_{is} > 0$. Because product i is not in the common set, Ω_{Cst}, the information of the price of commodity i is not considered in CUPI. Then, we have $\lambda_t = 1$ while $\lambda_s < 1$. Therefore, COLI by RW becomes becomes:

$$
\begin{aligned}
&\ln P_t - \ln P_s \\
&= \frac{1}{N^{C_{st}}} \sum_{i \in \Omega_{C_{st}}} \left((\ln p_{it} - \ln p_{is}) - \frac{1}{(1-\sigma)} \left(\ln w_{it}^{C_{st}} - \ln w_{is}^{C_{st}} \right) \right) + \frac{1}{(1-\sigma)} (\ln \lambda_s) && (39) \\
&> \frac{1}{N^{C_{st}}} \sum_{i \in \Omega_{C_{st}}} \left((\ln p_{it} - \ln p_{is}) - \frac{1}{(1-\sigma)} \ln w_{it}^{C_{st}} - \ln w_{is}^{C_{st}} \right) && (40) \\
&= \ln CCV && (41)
\end{aligned}
$$

Therefore, there is a negative variety effect. In other words, even if nobody wants to purchase commodity i in period t, the COLI increases because of the disappearance of the commodity. This contradiction occurs because RW assume that supply shocks cuase product entry and exit. Next, consider a case wherein the taste parameter, φ_{it}, becomes very small but remains positive. Then, the commodity's expenditure share becomes very small. Clearly, there were no variety effects. The problem here is that as φ_{it} approaches zero, $\ln w_{it}^{C}$ goes to minus infinity. When it reaches zero, CCV cannot be defined. In other words, there is discontinuity in the CCV at $\varphi_{it} = 0$. Note that such discontinuity does not occur in the following definition of the COLI as long as $\sigma > 1$.

$$COLI^{C}(s,t) = \left(\frac{\sum_{i=1}^{N_t} \left(\frac{p_{it}}{\varphi_{it}} \right)^{1-\sigma}}{\sum_{i=1}^{N_t} \left(\frac{p_{is}}{\varphi_{is}} \right)^{1-\sigma}} \right)^{\frac{1}{1-\sigma}} . \tag{42}$$

Finally, the index by RW or Feenstra (1994) is not generally transitive. As Abe and Rao (2020) show, as long as all commodities are common goods, the price index is transitive and not subject to chain drifts. However, when there are entries or (and) exits of commodities, the normalization condition, (36), becomes specific to the combination of times, s and t, which subjects the CCV and RW indexes to chain drifts.

4.2 Ordinal Utility

The cardinal COLI approach proposed by Redding and Weinstein (2020) can be expressed in a more general form as follows,

$$\begin{aligned} COLI\,(s,t,U)\_C &= \frac{E\left(\frac{p_{1t}}{\varphi_{1t}}, \frac{p_{2t}}{\varphi_{2t}}, \ldots \frac{p_{Nt}}{\varphi_{Nt}}, 1\right) \times U}{E\left(\frac{p_{1s}}{\varphi_{1s}}, \frac{p_{2s}}{\varphi_{2s}}, \ldots \frac{p_{Ns}}{\varphi_{Ns}}, 1\right) \times U}, \\ &= \frac{C\left(\frac{p_{1t}}{\varphi_{1t}}, \frac{p_{2t}}{\varphi_{2t}}, \ldots \frac{p_{Nt}}{\varphi_{Nt}}\right)}{C\left(\frac{p_{1s}}{\varphi_{1s}}, \frac{p_{2s}}{\varphi_{2s}}, \ldots \frac{p_{Ns}}{\varphi_{Ns}}\right)} \end{aligned} \tag{43}$$

When deriving (43) , we select the following quantity vectors, q_t and q_s with which the following equation holds:

$$U\left(\varphi_{1t}q_{1t}, \varphi_{2t}q_{2t}, \ldots\varphi_{Nt}q_{Nt}\right) = U\left(\varphi_{1s}q_{1s}, \varphi_{2s}q_{2s}, \ldots\varphi_{Ns}q_{Ns}\right) \tag{44}$$

.

A monotonic transformation of the utility function changes the values of the COLI because the levels of the two utility functions are equated. In this sense, cardinal COLI is not invariant to monotonic transformations of the utility functions.

The other COLI with heterogeneous preferences is called ordinal COLI, originally proposed by Balk (1989). The ordinal COLI is defined as,

$$COLI\left(s,t,U\right)\_B=\frac{E\left(\frac{p_{1t}}{\varphi_{1t}},\frac{p_{2t}}{\varphi_{2t}},...\frac{p_{Nt}}{\varphi_{Nt}},1\right)\times U\left(\varphi_{1t}q_{1f},\varphi_{2t}q_{2f},...\varphi_{Nt}q_{Nf}\right)}{E\left(\frac{p_{1s}}{\varphi_{1s}},\frac{p_{2s}}{\varphi_{2s}},...\frac{p_{Ns}}{\varphi_{Ns}},1\right)\times U\left(\varphi_{1s}q_{1f},\varphi_{2s}q_{2f},...\varphi_{Ns}q_{Nf}\right)},$$
$$=\frac{C\left(\frac{p_{1t}}{\varphi_{1t}},\frac{p_{2t}}{\varphi_{2t}},...\frac{p_{Nt}}{\varphi_{Nt}}\right)\times U\left(\varphi_{1t}q_{1f},\varphi_{2t}q_{2f},...\varphi_{Nt}q_{Nf}\right)}{C\left(\frac{p_{1s}}{\varphi_{1s}},\frac{p_{2s}}{\varphi_{2s}},...\frac{p_{Ns}}{\varphi_{Ns}}\right)\times U\left(\varphi_{1s}q_{1f},\varphi_{2s}q_{2f},...\varphi_{Ns}q_{Nf}\right)} \tag{45}$$

where $q_f=(q_{1f},q_{2f,}...,q_{Nf})$ is an exogenous reference quantity vector. In (45), we compare the utility levels that are indifferent to the exogenous quantity vector. Because this COLI is invariable to the monotonic transformation of utility functions, (45) is called ordinal COLI.

When utility function is CES, the ordinal COLI can be written as

$$COLI\left(s,t,U\right)\_B=\frac{\left(\sum_{i=1}^{N_t}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}\times\left(\sum_{i=1}^{N_t}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N_s}\left(\frac{p_{is}}{\varphi_{is}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}\times\left(\sum_{i=1}^{N_s}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}. \tag{46}$$

From (46), we can observe several notable characteristics of the ordinal COLI. First, if the product variety in the comparison period, the numerator of (46) is greater than that of the base periods, the denominator of (46), given the price levels, the COLI tends to be greater. This is in sharp contrast to the variety effects of Feenstra (1994), in which the COLI becomes smaller as the product variety increases. This contrast is due to the difference in the mechanisms behind the product turnover. According to Feenstra (1994), product turnover occurs because of supply shocks. When a product disappears, in Feenstra's model, the disappearance of a product is equivalent to a huge increase in the price, which increases COLI. In contrast, in ordinal COLI, a product disappears because we do not obtain utility from the product. In ordinal COLI, we evaluate the unit cost function at a utility level that is indifferent from the reference vector. Suppose that at time t, the number of non-zero elements of φ_{it} becomes greater than that in time s. Then, the utility level at time t that is indifferent from the reference vector tends to be greater than the utility level at time s. For example, consider an example of ice cream. Suppose that during summer, people obtain utilities from various ice creams whereas in winter, people obtain little utility from ice cream. Thus, in summer, to obtain a utility level that is indifferent from that in winter from ice creams, we need to purchase many more ice creams.

A potential problem in ordinal COLI is that we need to specify all the taste parameters to consturct the COLI. Balk (1989) points out that the actual calculation of ordinal COLI would be very difficult. However, as far as we use the CES utility function, the following proposition shows that we do not need to estimate the parameters, φ_{it}, when obtaining ordinal COLI. We can easily compute the COLI from prices, expenditure shares, and the exogenous reference vector as long as we know the value of the elasticity of substitution, σ.

Proposition 2 *When $\sigma\neq1$, the ordinal COLI can be written as*

$$COLI\left(s,t,U\right)\_B=\frac{\left(\sum_{i=1}^{N_t}\left(p_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N_s}\left(p_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \tag{47}$$

where w_{it} is the expenditure share of commodity i at time t.

Proof. See the Mathematical Appendix. ■

Note that when the elasticity of substitution, σ, goes to (positive) infinity, we obtain

$$\begin{aligned}\lim_{\sigma\to\infty}COLI\left(s,t,U\right)\_B&=\lim_{\sigma\to\infty}\frac{\left(\sum_{i=1}^{N_t}\left(p_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N_s}\left(p_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}\\&=\frac{\sum_{i=1}^{N_t}\left(p_{it}q_{if}\right)}{\sum_{i=1}^{N_s}\left(p_{is}q_{if}\right)},\end{aligned} \tag{48}$$

which is the standard fixed basket price index. It is worth noting that when the elasticity of substitution, σ, goes to infinity (positive), the RW index becomes Jevons index over the common goods as follows;

$$\begin{aligned}&\lim_{\sigma\to\infty}\frac{1}{N^{C_{st}}}\sum_{i\in\Omega_{C_{st}}}\left(\left(\ln p_{it}-\ln p_{is}\right)-\frac{1}{(1-\sigma)}\left(\ln w_{it}^{C_{st}}-\ln w_{is}^{C_{st}}\right)\right)-\frac{1}{(1-\sigma)}\left(\ln\lambda_t-\ln\lambda_s\right)\\&=\sum_{i\in\Omega_{C_{st}}}\left(\ln p_{it}-\ln p_{is}\right).\end{aligned} \tag{49}$$

Another notable difference between the cardinal COLI by RW and our ordinal COLI becomes clear when elasticity approaches unity. While the RW index as well as the variety effects index by Feenstra (1994) cannot be defined when the elasticity is unity, the ordinal COLI becomes the Cobb-Douglas index as follows;

$$\begin{aligned}\lim_{\sigma\to 1}COLI\left(s,t,U\right)\_B&=\lim_{\sigma\to 1}\frac{\left(\sum_{i=1}^{N_t}\left(p_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N_s}\left(p_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}\\&=\frac{\prod_{i=1}^{N_t}\left(w_{it}\right)^{-w_{it}}}{\prod_{i=1}^{N_s}\left(w_{is}\right)^{-w_{is}}}\times\frac{\prod_{i=1}^{N_t}\left(p_{it}q_{if}\right)^{w_{it}}}{\prod_{i=1}^{N_s}\left(p_{is}q_{if}\right)^{w_{is}}}.\end{aligned} \tag{50}$$

The first part of (50) is identical to the price index proposed by Lewbel (1989).[8] The ordinal COLI has several desirable characteristics. First, we do not need to assume constant elasticity across different times

[8] Lewbel (1989) derives the price index where the preferences are heterogeneous across regions. See page 314-315 in Lewbel (1989).

or regions. The elasticity, σ, can be heterogeneous across economic states. We can show that ordinal COLI is transitive, commensurable, and linearly homogeneous with respect to comparison prices.

4.3 Product Turnover in the Ordinal COLI

In (47), there are no Lambda effects or any other specific terms to capture product entry and exit. It is noteworthy that in (47), the denominator and numerator take the summation of different commodity sets. That is, ordinal COLI can be defined over different sets quite simply. More precisely, when the quantity of commodity i becomes zero, as long as $\sigma > 0$, we have

$$\lim_{w_{is}\to 0} w_{is}^{\frac{1}{\sigma}} = 0, \tag{51}$$
$$\lim_{w_{is}\to 0} w_{is}^{-w_{is}} = 1.$$

That is, the COLI of (47), includes the effects of product entry and exit caused by demand shocks. A product disappears when its taste parameters become zero. Unlike the index by Feenstra (1994) or RW, we do not discard the price information of commodities that are not in the set of common goods.

Another notable characteristic of ordinal COLI, (47), is that when the number of products increases at time t, the numerator has more factors when taking the summations, which leads to greater COLI in general. This is the opposite of the variety effect in Feenstra's COLI, (33) which decreases when variety increases. For example, during summer, as indicated in Figure 2, the nominal expenditure for ice cream was greater than that in winter. When deflated by Feenstra's COLI, the real expenditure exhibits greater seasonality, whereas deflated by the ordinal COLI, the seasonality becomes smaller. This sharp contrast arises because of the differences in the mechanism of seasonality. According to Feenstra (1994) and Redding and Weinstein (2020), seasonality in product variety is caused by supply shocks. In other words, consumers prefer a greater variety in winter. However, owing to exogenous supply shocks, consumers cannot enjoy a greater variety of ice creams during winter. In ordinal COLI, seasonality occurs because consumers change their the marginal utility from that of ice cream. During winter, consumers do not enjoy much utility from ice cream, as in summer, thus decreasing variety.

5 The Reference Vector

The reference vector plays a critical role in ordinal COLI, (47) because it enables us to compare the minimum expenditures between two different preferences. In the literature on social choices and welfare economics, the use of references to make inter personal comparisons has long been proposed and discussed.[9] A potentially serious problem in the reference vector is that the order of the utility levels might depend on the choice of

[9] See Fleuebaey (2009) and Bosmans et al. (2018), for example.

the reference vector. That is, when denoting the utility at the reference vector q_f under the preferences s and t by $U(q_f, s)$ and $U(q_f, t)$, $U(q_f, s) < U(q_f, t)$ for some q_f and $U(q'_f, s) > U(q'_f, t)$ for some q'_f, While there is no solid economic theory that justifies a particular reference vector for the ordinal COLI, we can show a few popular axioms in index number theory characterize a unique reference vector. Before going into details of the characterization, let us examine how different reference vectors lead to different index numbers.

To simplify the discussion, let us consider the case with constant variety. That is, we assume that the number of commodities is always N over time. Suppose we set $q_f = q_s$.Then, the index becomes

$$PI(s,t;q_f) = \frac{\left(\sum_{i=1}^{N}(p_{it}q_{is})^{\frac{\sigma-1}{\sigma}}(w_{it})^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N}(p_{is}q_{is})^{\frac{\sigma-1}{\sigma}}(w_{is})^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \text{ if } \sigma \neq 1 \tag{52}$$

$$= \frac{\prod_{i=1}^{N}(w_{it})^{-w_{it}} \times \prod_{i=1}^{N}(p_{it}q_{is})^{w_{it}}}{\prod_{i=1}^{N}(w_{is})^{-w_{is}} \times \prod_{i=1}^{N}(p_{is}q_{is})^{w_{is}}} \text{ if } \sigma = 1. \tag{53}$$

If σ goes to ∞, the index becomes the Laspeyres index,

$$PI(s,t;q_f) = \frac{\sum_{i=1}^{N}(p_{it}q_{is})}{\sum_{i=1}^{N}(p_{is}q_{is})}. \tag{54}$$

If we set $q_f = q_t$,the index becomes the Paasche index when elasticity is infinite,

$$PI(s,t;q_f) = \frac{\sum_{i=1}^{N}(p_{it}q_{it})}{\sum_{i=1}^{N}(p_{is}q_{it})}. \tag{55}$$

If we adopt the arithmetic mean of quantities at s and t, we obtain the Marshall-Edgeworth Index when $\sigma = \infty$,

$$PI(s,t;q_f) = \frac{\sum_{i=1}^{N} p_{it}\frac{(q_{is}+q_{it})}{2}}{\sum_{i=1}^{N} p_{is}\frac{(q_{is}+q_{it})}{2}}. \tag{56}$$

Finally, when using the geometric mean of the quantities, we obtain the Walsh index when $\sigma = \infty$,

$$PI(s,t;q_f) = \frac{\sum_{i=1}^{N} p_{it}\sqrt{q_{is}q_{it}}}{\sum_{i=1}^{N} p_{is}\sqrt{q_{is}q_{it}}}. \tag{57}$$

5.1 Characterization of the Reference Vector

Suppose we consider the following functional form for q_f^{st} when comparing time s and t,

$$q_f^{st} = \left(q_{1f}^{st}, q_{2f}^{st}, ..., q_{Nf}^{st}\right)$$
$$m_i^{st} : \mathbf{R}_{++}^M \rightarrow \mathbf{R}_{++}$$
$$q_{if}^{st} = m_i^{st}\left(q_{i1}, q_{i2}, .., q_{iM}\right) \text{ for } i = 1, ..., N$$
$$m_i^{st}\left(a, a, ..., a\right) = a \text{ for any } a \in \mathbf{R}_{++} \text{ for } i = 1, ..., N$$

where M is the number of regions or times where q_{ik} is the quantity of commodity i at time (location) k. Note that the domain of the reference function, m_i^{st}, is not restricted to the two comparing states, s and t as in the Laspeyres or Walsh indexes, but the domain contains all the economic states.

We denote $q_i = \left(q_{1i}, q_{i2}, .., q_{iM}\right)$. Then, the price index from states s to t can be written as,

$$PI\left(s, t; q_f^{st}\right) = \frac{\left(\sum_{i=1}^N \left(p_{it} m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}} \left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^N \left(p_{is} m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}} \left(w_{is}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \text{ if } \sigma \neq 1 \tag{58}$$

$$= \frac{\prod_{i=1}^N \left(w_{it}\right)^{-w_{it}} \times \prod_{i=1}^N \left(p_{it} m_i^{st}\left(q_i\right)\right)^{w_{it}}}{\prod_{i=1}^N \left(w_{is}\right)^{-w_{is}} \times \prod_{i=1}^N \left(p_{is} m_i^{st}\left(q_i\right)\right)^{w_{is}}} \text{ if } \sigma = 1 \tag{59}$$

Similarly, the price index from states t to s is given by

$$PI\left(t, s; q_f^{ts}\right) = \frac{\left(\sum_{i=1}^N \left(p_{is} m_i^{ts}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}} \left(w_{is}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^N \left(p_{it} m_i^{ts}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}} \left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \text{ if } \sigma \neq 1 \tag{60}$$

$$= \frac{\prod_{i=1}^N \left(w_{is}\right)^{-w_{is}} \times \prod_{i=1}^N \left(p_{is} m_i^{ts}\left(q_i\right)\right)^{w_{it}}}{\prod_{i=1}^N \left(w_{it}\right)^{-w_{it}} \times \prod_{i=1}^N \left(p_{it} m_i^{ts}\left(q_i\right)\right)^{w_{is}}} \text{ if } \sigma = 1. \tag{61}$$

We denote the quantity vector in which q_{is} and q_{it} are exchanged as $q_{i\_st} = \left(q_{i1}, ., q_{is}, ..., q_{it}, .., q_{iM}\right)$. Then, by construction, the following equation must hold for all $i = 1, 2, ..., N$ and $s, t = 1, 2, ..., M$,

$$m_i^{ts}\left(q_i\right) = m_i^{st}\left(q_{i\_st}\right). \tag{62}$$

Now, let us introduce the following three axioms, which have played major roles in index number theory since Fisher (1922):

Definition 3 *Symmetric function: A function $m_i : \mathbf{R}_{++}^M \to \mathbf{R}_{++}$ is symmetric if it is invariant to changes in order of the variables, that is, for any $s \neq t$,we have*

$$m_i\left(q_{i1}, ., q_{is}, ..., q_{it}, .., q_{iM}\right) = m_i\left(q_{i1}, ., q_{it}, ..., q_{is}, .., q_{iM}\right) \tag{63}$$

Definition 4 *State Reversal: for any $s, t,\ p_t, p_s, q_s,$and $q_t,$ we obtain*

$$PI\left(s, t; q_f^{st}\right) \times\ PI\left(t, s; q_f^{ts}\right) = 1 \tag{64}$$

Definition 5 *Transitivity: $PI\left(s, t; q_f^{st}\right)$ is transitive if for any $s, t, k,$ we always have*

$$PI\left(s, t; q_f^{st}\right) \times PI\left(t, k; q_f^{tk}\right) = PI\left(s, k; q_f^{sk}\right) \tag{65}$$

The above axioms impose a useful structure on the reference vector, m_i^{st}.Note that if both prices and quantities at s, t are identical, we get

$$PI\left(s, s; q_f\right) = \frac{\left(\sum_{i=1}^N \left(p_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}} \left(w_{is}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^N \left(p_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}} \left(w_{is}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} = 1. \tag{66}$$

That is, $PI\left(s, t; q_f^{st}\right)$ passes (weak) identity test. Therefore, if $PI\left(s, t; q_f^{st}\right)$ is transitive, $PI\left(s, t; q_f^{st}\right)$ passes the state reversal test.

Proposition 6 *Suppose $PI\left(s, t; q_f^{st}\right)$ is transitive. Then, for all $s, t = 1, 2, ..., M,$ we have*

$$m_i^{st}\left(q_i\right) = m_i\left(q_i\right), \tag{67}$$

and

$$m_i\left(q_i\right) = m_i\left(q_{i\_st}\right). \tag{68}$$

That is, all the reference vectors have identical symmetric functional forms across economic states.

Proof. For the proof, see Mathematical Appendix. ■

If the price index, $PI\left(s, t; q_{stf}\right)$, is transitive, the above proposition shows that, for any s, t, the reference vector must take the following form:

$$q_f^{st} = \left(m_1\left(q_1\right), m_2\left(q_2\right), .., m_N\left(q_N\right)\right) \text{ for all } s.t. \tag{69}$$

where $q_i = \left(q_{i1}, q_{i2}, ..., q_{iM}\right)$.

That is, the domain of the price index, $PI\left(s, t; q_f^{st}\right)$, is $p_s, p_t,\ q_s, q_t,$ and the quantities in all states. Thus, the index becomes multilateral. To be transitive, the reference vector needs to be common across all the comparisons. Thus, we can drop the upper script, st, from q_f^{st}. We denote the index as

$$PI(s,t;q_f) = PI(p_s,p_t,q) \tag{70}$$
$$\text{where } q = (q_1,...,q_N).$$

Before going into the main proposition, let us introduce the following Lemma;

Lemma 7 *Let* $q_1 = (q_{11}, q_{21}..., q_{N1}), q_2 = (q_{12}, q_{22}..., q_{N2}) \in \mathbf{R}_+^N$. *Suppose for all* $\lambda > 0$ *and* q_1, q_2, *the following equation holds:*

$$\frac{m_1(\lambda q_{11}, q_{21},..)}{m_1(q_{11}, q_{21},...)} = \frac{m_2(\lambda q_{12}, q_{22},..)}{m_2(q_{12}.q_{22},..)}. \tag{71}$$

Then, there exists a function, $f(\lambda)$, *that satisfies*

$$\frac{m_1(\lambda q_{11}, q_{21},..)}{m_1(q_{11}, q_{21},...)} = \frac{m_2(\lambda q_{12}, q_{22},..)}{m_2(q_{12}.q_{22},..)} = f(\lambda). \tag{72}$$

Proof. Set $(q_{11}, q_{21}..., q_{N1}) = (1, 1..., 1)$,then, for any $q_2 \in \mathbf{R}_+^N$, the following equation holds:

$$\frac{m_1(\lambda, 1..)}{m_1(1, 1...)} = \frac{m_2(\lambda q_{12}, q_{22},..)}{m_2(q_{12}.q_{22},..)}. \tag{73}$$

Similarly, set $(q_{12}, q_{22}..., q_{N2}) = (1, 1..., 1)$ .Then, for any $q_1 \in \mathbf{R}_+^N$, the following equation holds:

$$\frac{m_1(\lambda q_{11}, q_{21},..)}{m_1(q_{11}, q_{21},...)} = \frac{m_2(\lambda, 1, 1, ..)}{m_2(1.1..)} \tag{74}$$

Because for any $q_1, q_2 \in \mathbf{R}_+^N$, (74) must hold, we obtain

$$\begin{aligned} \frac{m_1(\lambda q_{11}, q_{21},..)}{m_1(q_{11}, q_{21},...)} &= \frac{m_2(\lambda q_{12}, q_{22},..)}{m_2(q_{12}.q_{22},..)} \\ &= \frac{m_1(\lambda, 1..)}{m_1(1, 1...)} \\ &= \frac{m_2(\lambda, 1, 1, ..)}{m_2(1.1..)}. \end{aligned} \tag{75}$$

We define

$$f(\lambda) = \frac{m_1(\lambda, 1..)}{m_1(1, 1...)}. \tag{76}$$

Then, we get

$$\frac{m_1(\lambda q_{11}, q_{21},..)}{m_1(q_{11}, q_{21},...)} = \frac{m_2(\lambda q_{12}, q_{22},..)}{m_2(q_{12}.q_{22},..)} = f(\lambda). \tag{77}$$

■

Following axiom is the last that enables us to characterize the reference vector.

Definition 8 *Invariant to proportional changes of a state*

Suppose in a state, j, all the quantities are multiplied by $\lambda > 0$, that is, the new quantity vector for state j becomes,

$$\widetilde{q}_j = \lambda \left(q_{1j}, q_{2j}, ..., q_{Nj} \right), \tag{78}$$

$$\widetilde{q} = \left(q_1, ..., q_{j-1}, \widetilde{q}_j, q_{j+1}, .., q_M \right), \tag{79}$$

then, $PI\left(p_s, p_t, q\right)$ is unchanged, that is,

$$PI\left(p_s, p_t, q\right) = PI\left(p_s, p_t, \widetilde{q}\right) \tag{80}$$

This axiom is regarded as one of the most important in international price comparisons. Without this property, the price index is significantly affected by the prices and quantities of large countries. We can also interpret the axiom as a requirement, so that the price index should not be affected by the total consumption or per capita consumption of a country is used. For temporal comparisons, the axiom implies that common seasonal fluctuations in quantities do not affect the price index.[10]

The following proposition shows the characterization of the reference vector for the ordinal COLI.

Proposition 9 *Suppose m_i is an increasing function of q_{im},or m_i is a continuous function. If $PI\left(p_s, p_t, q\right)$ is invariant to proportional changes of any one state and passes the transitivity test, m_i should have the following functional form,*

$$m_i = \prod_{m=1}^{M} \left(q_{im}\right)^{1/M}. \tag{81}$$

Proof. For the proof, see the Mathematical Appendix. ■

Proposition 9 enables us to uniquely determine the reference vector. In general, choosing the reference vector for evaluating the utility level across different individuals is a very difficult task. It is possible to characterize the reference vector because our purpose of conducting inter personal comparisons is to compare the minimum expenditures across individuals. Because the minimum expenditure is cardinal, we have a strong restriction on the comparisons, which makes a sharp contrast to the more general objectives such as welfare comparison among people. Note that if we restrict the domain of the reference vector to the quantities at states s and t, and drop the transitivity axiom, we obtain the Walsh Index when the elasticity of substitution is infinite, which is known to be a superlative index by Diewert (1976). That is, the ordinal COLI, (47), contain not only the Laspeyres, and the Paasche indexes, but also a superlative index such as the Walsh index as its special case.

[10] See Diewert (2001, p.207) for the discussion.

5.2 Reference Vector with Product Turnover

When there are product turnovers, or differences in the commodity sets between two different states, the quantities of the product that go out of the market, or that do not exist in the country are zero. Suppose there are M states or times. If one of the states or times contains zero quantity, the geometric formula of the reference vector, (81) , becomes zero even if all other states report a very large quantity. This implies that we lose much information about the commodity that has zero quantity in one of the states. To mitigate this problem, we must approximate (81). Therefore, we do not need to discard much information. Among the many possible approximations, our preferred method is to use a function known as a generalized mean of order r, as follows:

$$m_i = \left(\frac{1}{M} \sum_{m=1}^{M} \left(q_{im}^{z_i} \right) \right)^{1/z_i}, \tag{82}$$

where

$$z_i = \frac{1}{M} \sum_{m=1}^{M} I_{im} \left(q_{im} = 0 \right). \tag{83}$$

I_{im} is an indicator function that takes unity when quantity i at state m is zero. z_i is the ratio of observations of quantities that are zero. Then, as z_i approaches zero, (82) will converge to (81) , that is, we get

$$\lim_{z_i \to 0} \left(\frac{1}{M} \sum_{m=1}^{M} \left(q_{im}^{z_i} \right) \right)^{1/z_i} = \prod_{m=1}^{M} \left(q_{im} \right)^{1/M}.$$

5.3 Both Demand and Supply Shocks

When a product appears or disappears because of changes in the taste parameters, the ordinal COLI can automatically handle such turnover. However, if product turnover is caused by supply shocks, as in Feenstra (1994) or RW, we need to devise the COLI. Suppose that the preferences are fixed. In such cases, product turnover occurs purely because of changes in supply. Suppose p_{it} increases. If all other variables and parameters are unchanged, this will increase the quality (taste) adjusted price, p_{it}/φ_{it}. Given the taste parameters, an increase in price will increase the COLI. If p_{it} increases to infinity, as long as $\sigma > 1$, $\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}$ will converge to zero. That is,

$$\lim_{p_{it} \to \infty} \left(\frac{p_{it}}{\varphi_{it}} \right)^{1-\sigma} = \lim_{\varphi_{it} \to 0} \left(\frac{p_{it}}{\varphi_{it}} \right)^{1-\sigma} = 0.$$

Now, consider the following ordinal COLI,

$$COLI\left(s,t,U\right)\_B=\frac{\left(\sum_{i=1}^{N_t}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}\times\left(\sum_{i=1}^{N_t}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N_s}\left(\frac{p_{is}}{\varphi_{is}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}\times\left(\sum_{i=1}^{N_s}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}. \tag{84}$$

The increase in the price of good j affects the first part of the R.H.S as follows,

$$\lim_{p_{jt}\to\infty}\frac{\left(\sum_{i=1}^{N}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}}{\left(\sum_{i=1}^{N}\left(\frac{p_{is}}{\varphi_{is}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}}=\frac{\left(\sum_{i=1}^{j-1}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}+\left(\sum_{i=j+1}^{N}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}\right)^{\frac{1}{1-\sigma}}}{\left(\sum_{i=1}^{N}\left(\frac{p_{is}}{\varphi_{is}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}}. \tag{85}$$

On the other hand, the second part of the R.H.S. becomes,

$$\lim_{p_{jt}\to\infty}\frac{\left(\sum_{i=1}^{N}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}=\frac{\left(\sum_{i=1}^{N}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}},$$

$$\lim_{\varphi_{jt}\to 0}\frac{\left(\sum_{i=1}^{N}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}=\frac{\left(\sum_{i=1}^{j=1}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}+\sum_{i=j+1}^{N}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}.$$

Therefore, when we have a missing product caused by the supply factor, COLI becomes:

$$COLI\left(s,t,U\right)\_BS=\frac{\left(\sum_{i=1}^{j-1}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}+\left(\sum_{i=j+1}^{N}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}\right)^{\frac{1}{1-\sigma}}}{\left(\sum_{i=1}^{N}\left(\frac{p_{is}}{\varphi_{is}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}}\frac{\left(\sum_{i=1}^{N}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}. \tag{86}$$

If the disappearance is caused by the demand factor, the COLI becomes

$$\begin{aligned}&COLI\left(s,t,U\right)\_BD\\&=\frac{\left(\sum_{i=1}^{j-1}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}+\left(\sum_{i=j+1}^{N}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}\right)^{\frac{1}{1-\sigma}}}{\left(\sum_{i=1}^{N}\left(\frac{p_{is}}{\varphi_{is}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}}\times\frac{\left(\sum_{i=1}^{j=1}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}+\sum_{i=j+1}^{N}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}.\end{aligned} \tag{87}$$

The only difference between the two COLIs is in the utility levels of the reference quantity vector. To obtain $COLI\left(s,t,U\right)\_BS$, we need to know the value of φ_{it} even if we do not have information on prices at time t. To estimate $COLI\left(s,t,U\right)\_BD$, (47) can be used. In other words, (47) automatically includes the effects of product turnovers caused by preference changes.

Obviously, we obtain

$$
\begin{aligned}
&COLI\,(s,t,U)\,\_BS \\
&= COLI\,(s,t,U)\,\_BD \times \frac{\left(\sum_{i=1}^{N}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N}(\varphi_{is}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \times \frac{\left(\sum_{i=1}^{N}(\varphi_{is}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{j=1}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}+\sum_{i=j+1}^{N}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \\
&= COLI\,(s,t,U)\,\_BD \times \frac{\left(\sum_{i=1}^{N}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{j=1}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}+\sum_{i=j+1}^{N}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}
\end{aligned}
$$

A possible way to obtain φ_{it} is to assume $\varphi_{it}=\varphi_{is}$. If commodity is available at time s, we can easily obtain φ_{is}. The main steps are as follows. First, we find a commodity with positive quantities for both s and t. If there are several commodities that take positive values for all sample periods, we choose the product whose expenditure share is the largest. We denote such a commodity as commodity 1 and set $\varphi_{1t}=\varphi_{is}=1$.Then, from the first-order condition, we have

$$
\varphi_{is}=\left(\frac{p_{is}}{p_{1s}}\right)\left(\frac{w_{is}}{w_{1s}}\right)^{\frac{1}{\sigma-1}}.
$$

When we take the natural logarithms of both sides, we get

$$
\ln\varphi_{is}=\ln(p_{is})-\ln(p_{1s})+\frac{1}{\sigma-1}(\ln w_{is}-\ln w_{1s}). \tag{88}
$$

Then, set

$$
\ln\varphi_{is}=\ln\varphi_{it}.
$$

Then, compute

$$
\frac{\left(\sum_{i=1}^{N}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{j=1}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}+\sum_{i=j+1}^{N}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}.
$$

Finally, we obtain

$$
COLI\,(s,t,U)\,\_BS=\frac{\left(\sum_{i=1}^{N}(p_{it}q_{if})^{\frac{\sigma-1}{\sigma}}(w_{it})^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N}(p_{is}q_{if})^{\frac{\sigma-1}{\sigma}}(w_{is})^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}\times\frac{\left(\sum_{i=1}^{N}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{j=1}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}+\sum_{i=j+1}^{N}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}
$$

This is the COLI when product turnover occurs because of the supply factor.

Generally, if there are multiple disappearing commodities, we can use

$$COLI\left(s,t,U\right)\_BS=\frac{\left(\sum_{i\in I_t}\left(p_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_s}\left(p_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}\times\frac{\left(\sum_{i\in I_s}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_t}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}$$

When a new product, commodity $N+1$, appears at time t the opposite procedure can be conducted. If the newly appearing product is indexed as $N+1$,we need to estimate φ_{N+1s}.Using the same definition for commodity 1, we can obtain φ_{it} from the following first order condition,

$$\varphi_{N+1t}=\left(\frac{p_{N+1t}}{p_{1t}}\right)\left(\frac{w_{N+1t}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}.$$

Take the natural logarithms of both sides, we get

$$\ln\varphi_{N+1t}=\ln\left(p_{N+1}\right)-\ln\left(p_{1t}\right)+\frac{1}{\sigma-1}\left(\ln w_{N+1t}-\ln w_{1t}\right).$$

Then, by setting

$$\ln\varphi_{N+1s}=\ln\varphi_{N+1t},$$

we obtain φ_{it}.

The COLI is

$$COLI\left(s,t,U\right)\_BS=\frac{\left(\sum_{i=1}^{N}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}}{\left(\sum_{i=1}^{N+1}\left(\frac{p_{is}}{\varphi_{is}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}}\times\frac{\left(\sum_{i=1}^{N+1}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N+1}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}. \tag{89}$$

If there are multiple appearing goods, COLI_BS becomes

$$COLI\left(s,t,U\right)\_BS \tag{90}$$

$$= \frac{\left(\sum_{i\in I_t}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}}{\left(\sum_{i\in I_s}\left(\frac{p_{is}}{\varphi_{is}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}} \times \frac{\left(\sum_{i\in I_t}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_s}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}$$

$$\times \frac{\left(\sum_{i\in I_s}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_t}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \times \frac{\left(\sum_{i\in I_t}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_t}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \tag{91}$$

$$= COLI\left(s,t,U\right)\_BD \times \frac{\left(\sum_{i\in I_s}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_t}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \times \frac{\left(\sum_{i\in I_t}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_t}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \tag{92}$$

$$= \frac{\left(\sum_{i\in I_t}\left(p_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_s}\left(p_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \times \frac{\left(\sum_{i\in I_s}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_t}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \tag{93}$$

If there are both multiple appearing and disappearing products, we can obtain

$$COLI\left(s,t,U\right)\_BS = \frac{\left(\sum_{i\in I_t}\left(p_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_s}\left(p_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}$$
$$\times \frac{\left(\sum_{i\in I_s}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_t}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \times \frac{\left(\sum_{i\in I_s}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_t}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}, \tag{94}$$

where I_k is the set of commodities whose quantities are positive at time k.

Note that

$$\frac{\left(\sum_{i\in I_s}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_t}\left(\varphi_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \times \frac{\left(\sum_{i\in I_s}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_t}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \tag{95}$$

corresponds with the Lambda ratio in Feenstra (1994). This captures the hypothetical impacts of product entry and exit on the utility levels. In this paper, we employ a CES utility function with variable preferences, whereas Fenestra (1994) assumes that preference parameters are fixed over time. Moreover, Feenstra (1994) uses the Sato-Vartia index as the COLI, which can be justified when the observed data are consistent with the consumer's optimization behavior. In this study, we assume a variable preference model in which we can always regard the observed data to be consistent with consumer's optimization behavior. (94) is different from the CUPI in RW in that we adopt the ordinal utilities while the CUPI relies on the cardinal utility, and, therefore on the selection of the normalization of the parameters. Unfortunately, index number formula (94) is not transitive. Violation of transitivity occurs because of the imputation of the missing taste parameters. A possible way to maintain transitivity is to infer the taste parameters from all time periods. More specifically, using commodity 1 as the benchmark such that $\varphi_{1t} = 1$ for all t, we can obtain φ_{N+1t} from the following first order condition:

$$\varphi_{N+1t} = \left(\frac{p_{N+1t}}{p_{1t}}\right)\left(\frac{w_{N+1t}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}. \tag{96}$$

We pool the above equations from all time periods, $0 \leq t \leq M$ when the product is not missing. Then, taking the natural logarithms of both sides and the simple arithmetic mean, we obtain

$$\ln \varphi_{N+1t} = \frac{1}{\#(T_c)} \sum_{s \in T_C} \left[\ln(p_{N+1s}) - \ln(p_{1s}) + \frac{1}{\sigma - 1}(\ln w_{N+1s} - \ln w_{1s})\right]. \tag{97}$$

The index with the above imputed taste parameters should be transitive.

6 Identification of Demand and Supply Shocks

Usually, it is very difficult to identify whether supply or demand factors cause product turnover. A simple and practical method when there is no information on the extent of demand and supply shock is to take the geometric mean of the $COLI(s,t,U)\_BS$ and $COLI(s,t,U)\_BD$, that is,

$$\frac{\left(\sum_{i\in I_t}(p_{it}q_{if})^{\frac{\sigma-1}{\sigma}}(w_{it})^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_s}(p_{is}q_{if})^{\frac{\sigma-1}{\sigma}}(w_{is})^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \times \left(\frac{\left(\sum_{i\in I_s}(\varphi_{is}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_t}(\varphi_{is}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \times \frac{\left(\sum_{i\in I_s}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i\in I_t}(\varphi_{it}q_{if})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}\right)^{1/2}. \tag{98}$$

This index is easy to construct without complex nonlinear estimation. But it is based on a strong assumption that demand and supply shocks are the same. In the absence of such information, in this section, we show how to estimate the degree of demand and supply shocks.

First, rewrite the first- order condition (88) as follows,

$$(\ln p_{it} - \ln p_{is}) - (\ln P_t - \ln P_s) + \frac{1}{(\sigma-1)}\ln(w_{it} - \ln w_{is}) = \ln\varphi_{it} - \ln\varphi_{it}. \tag{99}$$

$$\Delta \ln p_{it} = \Delta \ln P_t - \frac{1}{(\sigma-1)}\Delta \ln w_{it} + \Delta \ln \varphi_{it} \tag{100}$$

where $\Delta \ln x_{it} = \ln x_{it} - \ln x_{is}$. $\Delta \ln \varphi_{it}$ can be interpreted as the demand shock between times s and t in commodity i. Following Feenstra (1994), we can also use the following simple supply function,

$$\Delta \ln q_{it} = \omega \Delta \ln p_{it} + C_t + \Delta \ln \delta_{it}, \omega > 0 \tag{101}$$

while C_t is the category specific time effects that affect the supply of quantities. $C_t + \Delta \ln \delta_{it}$ can be interpreted as the supply shocks between times s and t. By employing the estimation strategy of Feenstra(1994) or other related papers such as Hottman et al. (2016), we can estimate σ and ω.

In practice, we estimate σ and ω by applying nonlinear GMM using the following three moment conditions,

$$\frac{1}{N^{C_{st}}} \sum_{i \in C_{st}} z_{i,t} \left(\Delta\Delta \ln q_{it} + \sigma \Delta\Delta \ln p_{it} \right) = 0,$$

$$\frac{1}{N^{C_{st}}} \sum_{i \in C_{st}} z_{i,t} \left(\Delta\Delta \ln q_{it} - \omega \Delta\Delta \ln p_{it} \right) = 0,$$

$$\frac{1}{N^{C_{st}}} \sum_{i \in C_{st}} \left(\Delta\Delta \ln q_{it} + \sigma \Delta\Delta \ln p_{it} \right) \left(\Delta\Delta \ln q_{it} - \omega \Delta\Delta \ln p_{it} \right) = 0,$$

where,

$$\Delta\Delta \ln q_{it} = \Delta \ln q_{it} - \Delta \ln q_{1t},$$

$$\Delta\Delta \ln p_{it} = \Delta \ln p_{it} - \Delta \ln q_{1t},$$

and $\sigma > 0$ and $\omega > 0$. For the instrumental variable z_{it}, we use the $\Delta\Delta \ln p_{i,t-2}$. The reference item for calculating the relative price change and relative quantity change was the product with the largest sales share among the products that were continuously sold during the entire period. Estimation results based on this method were $\sigma = 8.64$ and $\omega = 14.98$.

Once we obtain the two consistent estimates of the elasticities, by putting the observed information into (100) and (101) , it is possible to obtain $\Delta \ln \varphi_{it}$ and $C_t + \Delta \ln \delta_{it}$, which are demand and supply shocks for product i between time periods (or locations) s and t. Next, we aggregate the commodity level demand and supply shocks to create category level shocks. Although there are several ways to aggregate the shocks, we prefer Tornqvist Index. Other formulas such as the Sato-Vartia, or Fisher do not create much differences. We do not use our ordinal COLI because the COLI assumes that all the shocks are caused by demand side. One potential problem in the above procedure is that we need to take the time (locations) differences to obtain the shocks. That is, we must drop the information of the product entry and exit when estimating the demand and supply shocks. Therefore, we need to assume that the demand and supply shocks that create product entry and exit are closely correlated with the demand and supply shocks observed in the continuing goods.

7 Empirical Examples

7.1 Weekly Scanner Data

Weekly scanner data on processed food and daily necessities are being increasingly used as official data in many countries. While scanner data contain very rich information about transactions, they are also known to exhibit a high level of product turnover. In this study, we use Japanese store-level weekly scanner data,

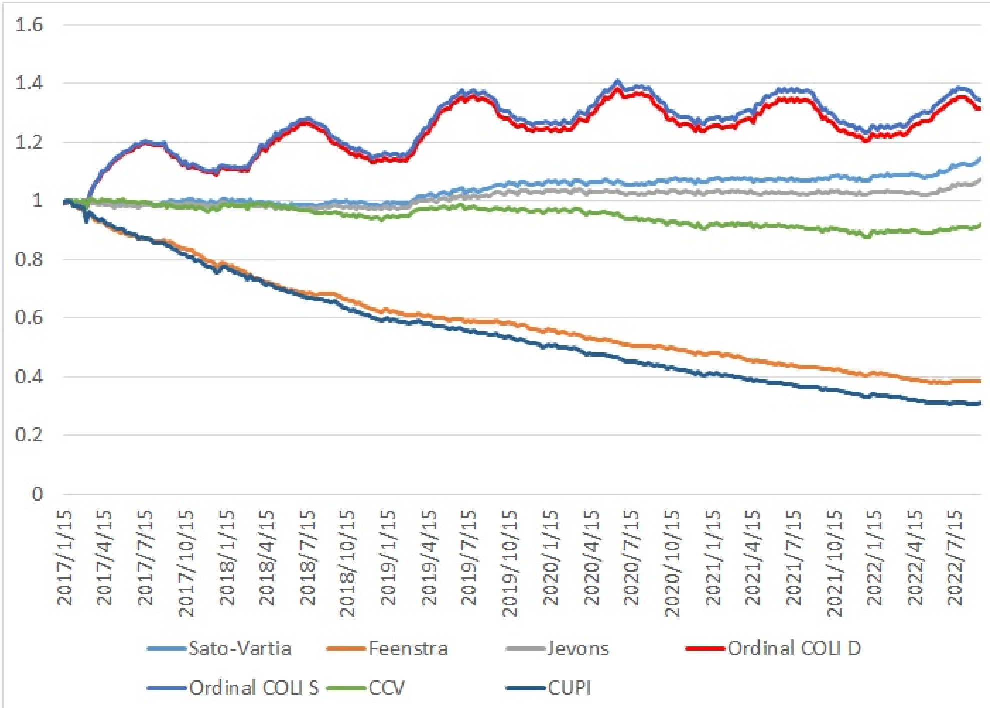

Figure 4: Ordinal COLI and Other Indexes

known as the SRI, collected by INTAGE Inc. Sales records in the dataset cover sales of processed foods, daily necessities, cosmetics, and drugs that have a commodity code known as the Japanese Article Number (JAN). In particular, we selected ice cream and snacks for this study

Figure 4 shows our ordinal COLI as well as the Feenstra index and the CUPI by Redding and Weinstein. For Feenstra index and CUPI, we show direct as well as chained indices. We show only direct Ordinal COLI because Ordinal COLI is transitive. Figure 4 indicates that both Feenstra's index and the CUPI proposed by Redding and Weinstein (2020) exhibit serious chain drifts. Second, ordinal COLI is subject to strong seasonality. During summer, ordinal COLI tends to increase, reflecting an increase in product varieties. These seasonal movements create the general movement of the "real expenditure," the nominal expenditure deflated by the price index, which is stable.

Figures 5 and **??** show differences in the real expenditures cased by the chain and direct price indexes, respectively. As is clear from the figures, the real ice cream expenditure deflated by the ordinal COLI exhibits smaller seasonal movements compared to nominal expenditure or real expenditures using other price index series as deflators.

7.2 Monthly Official CPI: Fruits

The official Japanese consumer price index (CPI) is constructed from the Familiy Expenditure Survey (FES) and the elementary level CPI. In this study, we use the official statistics of monthly expenditure of Japanese families by the FES and monthly item-level prices for each prefecture in Japan.

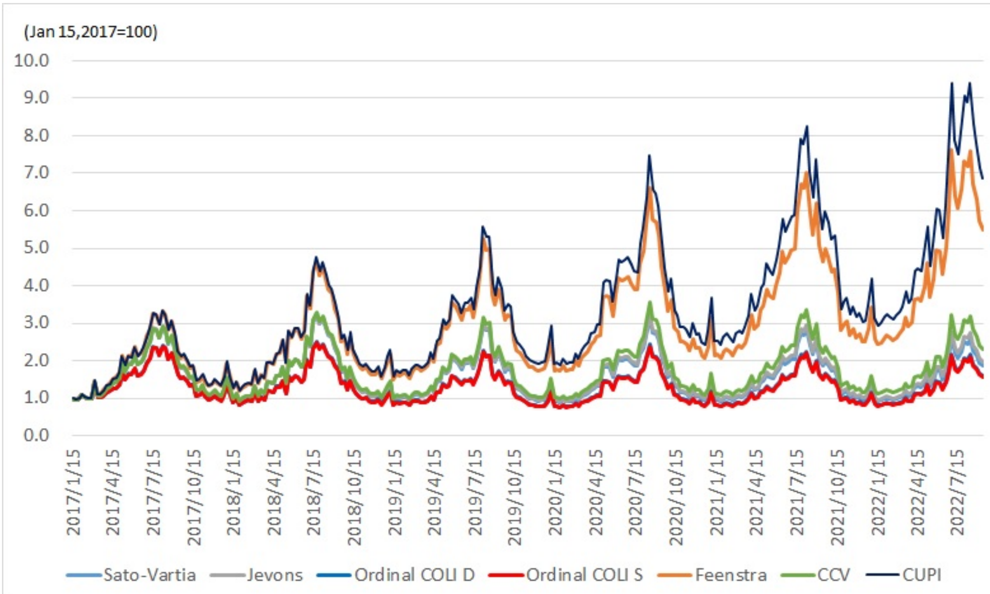

Figure 5: Real Expenditures by Different Chain Price Indexes

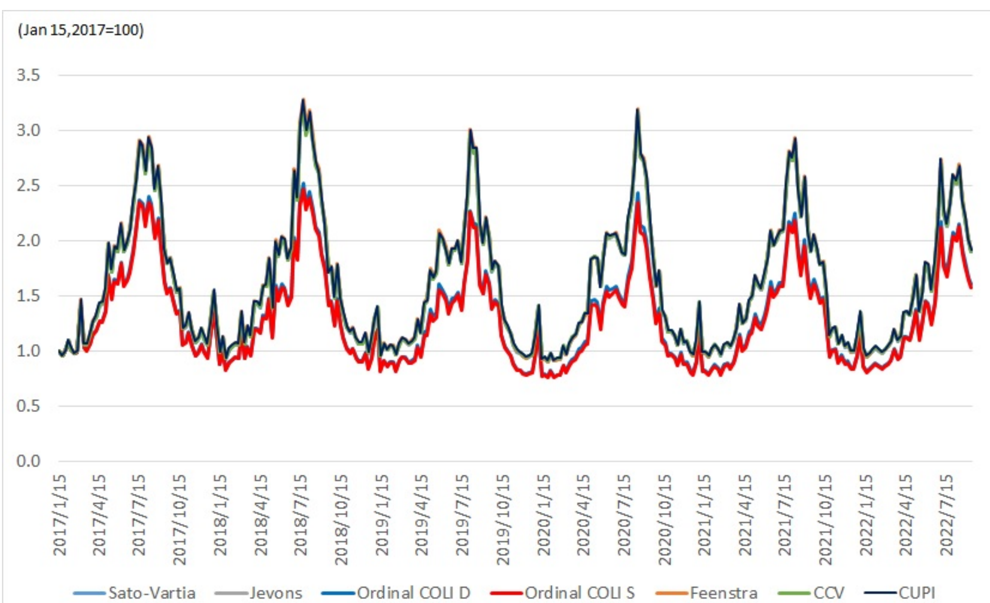

Figure 6: Real Expenditures by Different Direct Price Indexes

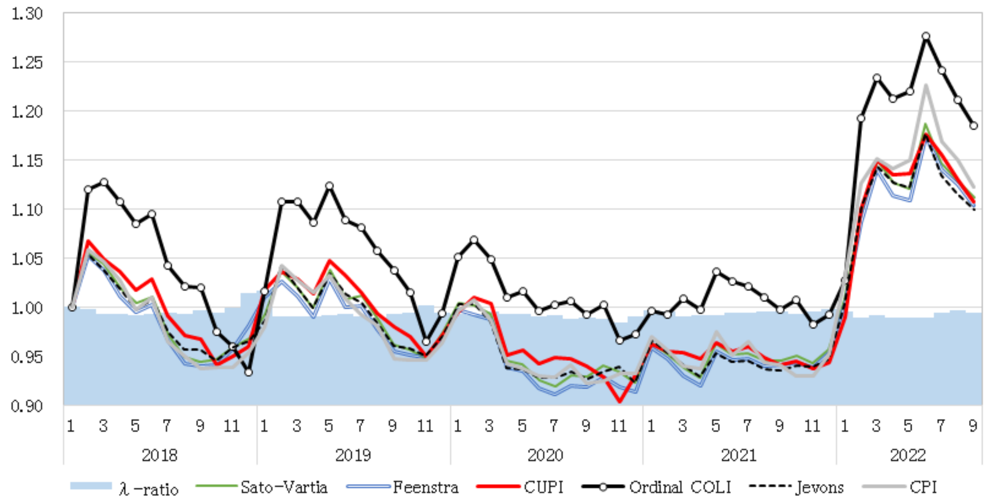

Figure 7: Price Index of Orange

Figure 7 shows the direct price index of orange while Figure 8 shows differences between the Direct and the Chain index (Direct Index - Chain Index). Because oranges occasionally become unavailable, the Jevons Index is not transitive. As in the ice cream, the ordinal COLI exhibits greater seasonality than other indices.

Figure 9 shows the differences in real expenditure in oranges with various index numbers. Although the differences are not large, the real expenditures obtained by ordinal COLI exhibit smaller volatility than other real expenditures. In the price survey we used, the change in variety comes from prefecture-level differences in the availability of oranges. That is, the variety effects from various “different” oranges such as organic orange; orange imported from abroad is not reflected in the figure, which mitiates the degree of the chain drift to a great extent. However, we can still observe the effects of product turnover on volatility in real expenditure

7.3 International Comparisons Program (ICP) data for PPP

The final empirical application of our index is an application to the purchasing power parity (PPP). The International Comparison Program (ICP) of the World Bank compiles price levels in different countries on a global scale. The ICP used the GEKS index number formula, which is a transitive index build using bilateral Fisher Indexes. Table 1 presents PPPs of currencies computed using different index number formulae. Because the ICP makes use of data at the category level in its computation of PPPs, the elasticity of substitution, such as the substitution between fruits and educational services, is supposed to be small. However, when the elasticity of substitution is smaller than unity, neither Feenstra’s index nor the CUPI are not well-defined. As is clear from Table 1, CUPI and Feenstra’s index exhibit very different values from the GEKS based on the Fisher index used in PPP computation in the the ICP . Compared to CUPI and

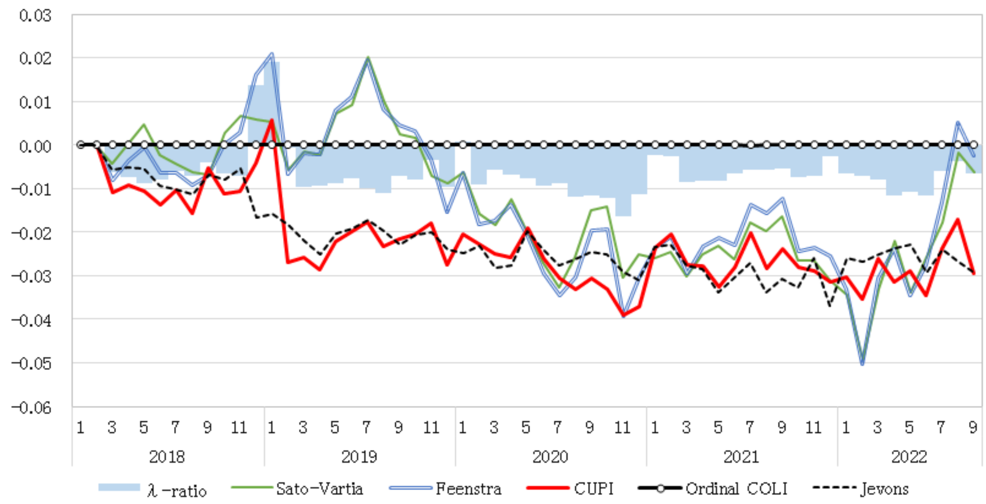

Figure 8: Gap between Direct and Chain Indices: Oranges

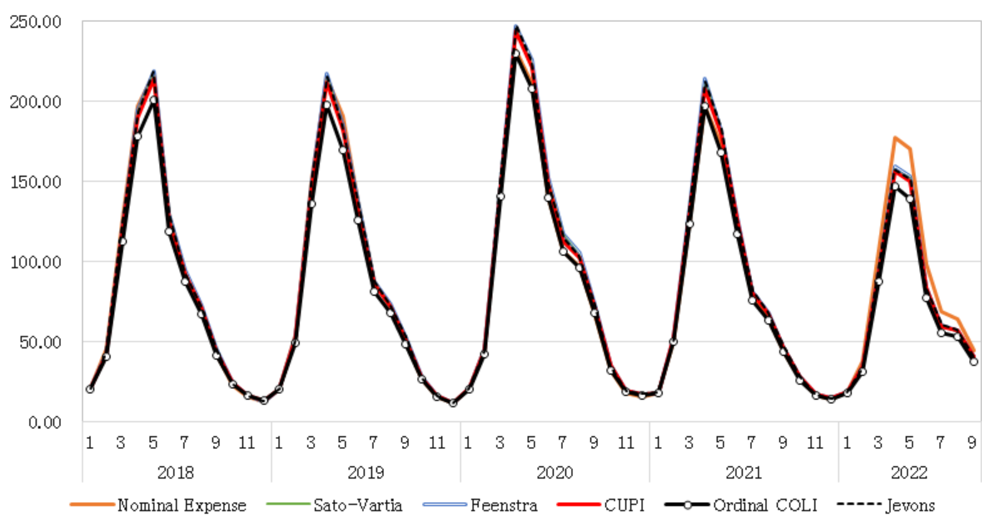

Figure 9: Real Expenditures on Orange by Direct Index

Table 1: Purchasing Power Parties Based on ICP 2017

Country	Ordinal COLI (0.9)	Ordinal COLI (1.1)	Feenstra (1.1)	CUPI (1.1)	Ordinal COLI (1)	Fisher	GEKS Fisher	Geary Khamis
Cameroon	266.52	263.79	238.32	73.04	265.11	239.77	240.91	213.56
Kenya	40.64	39.62	40.84	246.39	40.01	42.32	43.18	39.10
China	5.26	.13	8.13	83.03	5.20	4.03	4.21	4.02
Myanmar	329.48	333.29	516.23	0.09	331.48	400.22	397.09	361.53
Australia	1.70	1.72	1.60	6.68	1.71	1.51	1.51	1.48
Switzerland	1.22	1.24	0.99	4.13	1.23	1.27	1.33	1.27
Germany	0.91	0.89	0.65	16.18	0.90	0.80	0.80	0.78
United Kingdom	0.93	0.92	0.76	2.53	0.92	0.82	0.77	0.76
Japan	145.93	138.79	96.25	18.29	141.95	108.75	118.18	111.66
New Zealand	2.12	2.03	1.67	38.74	2.07	1.61	1.57	1.53
United States	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Peru	2.75	2.57	1.93	10.27	2.65	1.85	1.95	1.88
Saudi Arabia	2.44	2.40	2.26	6.71	2.42	1.65	1.71	1.54
Sudan (WAS)	6.26	6.08	6.42	1.50	6.17	5.33	5.48	5.23

Note:

The numbers in parentheses are the elasticity of substitution. The values are normalized so that the PPP for the United States is unity.

Feenstra's indices, ordinal COLI is much closer to GEKS. In some countries such as Japan, China, and Saudi Arabia, the discrepancies between GEKS and the Ordinal COLI are relatively greater, while differences in Germany and the United Kingdom are relatively small. We suspect that the differences are related to the degree of heterogeneity in preferences.

8 Conclusion

In this paper we have presented a framework for measuring changes in prices and real expenditures in the presence of product turnover, a phenomenon observed in frequency scanner data as well as data at higher levels of aggregation. The current approach to deal with differing commodity sets and the variety effects is the cost of living index proposed by Feenstra(1994) with costant elasticity substitution preferences. We show that the Feenstra index exhibits significant chain drift which in turn makes it difficult to make comparisons between two remote periods. Further, the Feenstra index can be defined only when the elsasticity of substitution is less than or equal to unity, which makes the index in application for aggregation at category or higher levels with substitution elasticities less than unity. We also discuss the recently proposed unified cost-of-living index proposed by Redding and Weinstein (2020) as a possible way of handling product turnover. The Redding and Wenstein approach uses hetrogeneous preferences and under the assunption of cardinal utility which is somewhat restrictive. Further, it is shown that their index is also subject to chain drift because product turn over, as in Feenstra (1994), is assumed to be purely due to supply shocks.

In our new approach, we replace cardinal utility approach of Redding and Weinstein with ordinal utility, and also allow the possibility for product turn over to be a result of supply as well as demand shocks. Based on Balk (1989), we advocate the use of reference quantity approach to multilateral price comparisons under heterogeneous preferences. Our approach does not require any normalization procedure similar to that necessary in the Redding and Wenstein (2020) approach. This approach is equally applicable at different levels of aggregation and can be applied even when elasticity of substitution is less than or equal to 1. In order to implement our approach, it is necessary to identify a suitable reference quantity vector for measuring the cost of living index under heterogenour preferenes. Using three standard axioms of index number theory, we characterise the reference vector which is subsequently used in all the ampirical applications. The paper has also developed a procedure to decompose the effects of demand and supply shocks on the cost of living index. Further, our index can also be applied to price data at different levels of aggregation: ranging from commodity-level scanner data to international price comparisons above the elementary level. In the empirical section, we have shown that our index has superior performance when applied to weekly scanner data. For example, in the case of ice cream for which there are strong seasonal trends in purchasing patterns and varieties, our index goes up during the summer, which makes the “real expenditures” on ice cream much smoother than the nominal expenditures, or other real expenditures deflated by Fisher’s index, Feenstra’s index, and the CUPI (CES-Unified Index) by Redding and Weinstein (2020). Furthermore, the CUPI and Feenstra’s index show a strongly negative drift, while our index is transitive. In the empirical section we have also demonstrated the applicability of our approach to national-level data as well as international comparison data are also presented. The results presented clearly show that our approach eliminates the strong negative drift associated with the Feenstra (1994) and the Redding and Weinstein (2020) indexes when applied to scanner data and that it leads to more meaningful price and real expenditure comparisons when applied to data from the International Comparison Program.

9 References

Abe, N and Rao, D.S.P (2020) ”Generalized Logarithmic Index Numbers with Demand Shocks: Bridging the Gap between Theory and Practice,” RCESR Discussion Paper Series DP20-1, Research Center for Economic and Social Risks, Institute of Economic Research, Hitotsubashi University.

Aczel, J. (1987) A Short Course on Functional Equations: Based Upon Recent Applications to the Social and Behavioral Sciences, D Reidel Publishing company, Dordrecht

Adelman, I and Z. Griliches (1961) On an Index of Quality Change, Journal of the American Statistical Association, 56:295, 535-548, DOI: 10.1080/01621459.1961.10480643

Balk, B.M. (1989) “Changing Consumer Preferences and the Cost-of-Living Index: Theory and Nonparametric Expressions,” Journal of ErZeitschrift fur National Ekonomie, 50-2, pp. 157-169.

Balk, B.M. (2000) “On Curing the CPI’s Substitution and New Goods Bias,” Research paper No. 0005 (Statistics Netherlands, Voorburg).

Bosmans, K., Decancq, K. and Ooghe, E. (2018),“ Who's afraid of aggregating money metrics?”, Theoretical Economics, 13: 467-484. https://doi.org/10.3982/TE2825

Court, Andrew T. (1939), “Hedonic Price Indexes with Automotive Examples”, in: The Dynamics of Automobile Demand, New York, NY: General Motors Corporation, pp. 99-117.

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.

Diewert, E.W. (1976), ”Exact and Superlative Index Numbers”, Journal of Econometrics 4: 115-145.

Diewert, W.E. (2001), “The Consumer Price Index and index number purpose”, Journal of Economic and Social Measurement 27 (2001) 167–248.

Diewert, W.E. (2018), “Scanner Data, Elementary Price Indexes and the Chain Drift Problem”, Discussion Paper 18-06, Vancouver School f Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1L4.

Diewert, W.E (2020), “The Chain Drift Problem and Multilateral Indexes”, Chapter 7, CPI Manual, ILO-ECE, Geneva. Forthcoming.

Feenstra, R.C. (1994), “New Product Varieties and the Measurement of International Prices”, American Economic Review 84, 157-177.

Fisher, I. (1911). The Purchasing Power of Money, London: Macmillan.

Fisher, I. (1922). The Making of Index Numbers, Houghton Mifflin Company, New York.

Fleurbaey, M. (2009) “Beyond GDP: The Quest for a Measure of Social welfare,” Journal of Economic Literature , Vol. 47, No. 4, pp. 1029-1075.

Lewbel, A. (1989). Identification and Estimation of Equivalence Scales under Weak Separability. The Review of Economic Studies, 56(2), 311–316. https://doi.org/10.2307/2297464

Luce, R.D. (1964) ”A generalization of a theorem of dimensional analysis,”Journal of Mathematical Psychology,Volume 1, Issue 2, 278-284.

Marshall, A. (1887), Contemporary Review, March.

Neary, J.P. (2004) ”Rationalizing the Penn World Table: True Multilateral Indices for International Comparisons of Real Income.” American Economic Review, 94 (5): 1411-1428.

Redding, S. J. and D. E. Weinstein (2020), “Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES Preferences”, Quarterly Joural of Economics, 503–560. doi:10.1093/qje/qjz031

Stone, R. (1956), Quantity and Price Indexes in National Accounts, Paris, France: Organization for European Economic Cooperation.

Triplett, J. (2004), “Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products”, OECD Science, Technology and Industry Working Papers, 2004/9, OECD Publishing.

World Bank (2020), Report Purchasing Power Parities and the Size of World Economies: Results from

the 2017 International Comparison Program, Washington DC.

A Mathematical Appendix

A.1 Proof for Proposition 2

Proof. If $\sigma \neq 1$, we obtain

$$COLI\left(s,t,U\right)\_B = \frac{\left(\sum_{i=1}^{N_t}\left(p_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N_s}\left(p_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \tag{A.1}$$

where w_{it} is an expenditure share of commodity i at state t. One of the first-order conditions for consumers' utility maximization is as follows;

$$\varphi_{it} = \left(\frac{p_{it}}{p_{1t}}\right)\left(\frac{w_{it}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}\varphi_{1t}. \tag{A.2}$$

Because $COLI\left(s,t,U\right)\_B$ is the homogeneous of degree zero with respect to $w_{it}\left(i=1,2,..N_t\right)$, we can set $\varphi_{1t}=1$ without it affecting $COLI\left(s,t,U\right)\_B$. Substituting $\varphi_{1t}=1$ to the first-order condition gives us

$$\varphi_{it} = \left(\frac{p_{it}}{p_{1t}}\right)\left(\frac{w_{it}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}. \tag{A.3}$$

A tedious calculation leads us to

$$\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma} = \frac{w_{it}}{p_{1t}^{\sigma-1}w_{1t}}. \tag{A.4}$$

We can also obtain the following relationships easily,

$$P_t = \left(\sum_{i=1}^{N}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}} = p_{1t}w_{1t}^{\frac{1}{\sigma-1}} \tag{A.5}$$

$$\varphi_{it}q_{if} = \frac{1}{p_{1t}}\left(\frac{1}{w_{1t}}\right)^{\frac{1}{\sigma-1}}\left(p_{it}q_{if}\right)\left(w_{it}\right)^{\frac{1}{\sigma-1}} \tag{A.6}$$

$$\begin{aligned}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}} &= \left(\frac{1}{p_{1t}}\right)^{\frac{\sigma-1}{\sigma}}\left(\frac{1}{w_{1t}}\right)^{\frac{1}{\sigma}}\left(\left(p_{it}q_{if}\right)\left(w_{it}\right)^{\frac{1}{\sigma-1}}\right)^{\frac{\sigma-1}{\sigma}} \\ &= \left(\frac{1}{p_{1t}}\right)^{\frac{\sigma-1}{\sigma}}\left(\frac{1}{w_{1t}}\right)^{\frac{1}{\sigma}}\left(\left(p_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)\end{aligned} \tag{A.7}$$

$$\left(\sum_{i=1}^{N_t}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} = \frac{1}{p_{1t}}\left(\frac{1}{w_{1t}}\right)^{\frac{1}{\sigma-1}}\left(\sum_{i=1}^{N_t}\left(p_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{A.8}$$

Therefore, the numerator of the $COLI\left(s,t,U\right)\_B$ can be written as

$$\begin{aligned}
&\left(\sum_{i=1}^{N_t}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}\times\left(\sum_{i=1}^{N_t}\left(\varphi_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}\\
&=p_{1t}w_{1t}^{\frac{1}{\sigma-1}}\times\frac{1}{p_{1t}}\left(\frac{1}{w_{1t}}\right)^{\frac{1}{\sigma-1}}\left(\sum_{i=1}^{N_t}\left(p_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}\\
&=\left(\sum_{i=1}^{N_t}\left(p_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \qquad \text{(A.9)}
\end{aligned}$$

A smilar calculation for the denominator gives us the following expression,

$$COLI\left(s,t,U\right)\_B=\frac{\left(\sum_{i=1}^{N_t}\left(p_{it}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N_t}\left(p_{is}q_{if}\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}. \qquad \text{(A.10)}$$

■

A.2 Proof for Proposition 6

Proof. Assume $\sigma\neq 1$. Then, by state reversal,

$$\begin{aligned}
&\left(PI\left(s,t;q_f^{ts}\right)\times PI\left(t,s;q_f^{ts}\right)\right)^{\frac{\sigma-1}{\sigma}}\\
&=\frac{\left(\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)}{\left(\sum_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)}\\
&\times\frac{\left(\sum_{i=1}^{N}\left(p_{is}m_i^{ts}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)}{\left(\sum_{i=1}^{N}\left(p_{it}m_i^{ts}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)}\\
&=1. \qquad \text{(A.11)}
\end{aligned}$$

Using $m_i^{ts}\left(q_i\right)=m_i^{st}\left(q_i\_st\right)$, we get

$$\begin{aligned}
&\frac{\left(\sum_{i=1}^{N}\left(p_{is}m_i^{ts}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)}{\left(\sum_{i=1}^{N}\left(p_{it}m_i^{ts}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)}\\
&=\frac{\left(\sum_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_{i_st}\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)}{\left(\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_{i_st}\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)}. \qquad \text{(A.12)}
\end{aligned}$$

Therefore, we obtain

$$\frac{\left(\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)}{\left(\sum_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)} \times \frac{\sum_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_{i\_st}\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_{i\_st}\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}} = 1 \tag{A.13}$$

Suppose we have

$$m_i^{st}\left(q_i\right) \neq m_i^{st}\left(q_i\_st\right).$$

Then, there exists $X \neq 1$ such that,

$$\frac{\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_{i\_st}\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}} = X \neq 1. \tag{A.14}$$

From $(A.13)$,

$$\frac{\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}} = \frac{\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_i\_st\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_i\_st\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}}. \tag{A.15}$$

Therefore,

$$\frac{\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_i\_st\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}} = \frac{\sum_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_i\_st\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}} = X \tag{A.16}$$

However, by choosing p_s,it is possible to make

$$\frac{\sum_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_i\_st\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}} \neq X.$$

This is a contradiction. Therefore, the following equation must always hold.

$$m_i^{st}\left(q_i\right) = m_i^{st}\left(q_{i\text{-}st}\right). \tag{A.17}$$

Because the index is transitive, we get

$$
\begin{aligned}
& PI\left(s,t;q_f\right)^{\frac{\sigma-1}{\sigma}} \times PI\left(t,k;q_f^{tk}\right)^{\frac{\sigma-1}{\sigma}} \\
& = \frac{\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}} \\
& \times \frac{\sum_{i=1}^{N}\left(p_{ik}m_i^{tk}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{ik}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{it}m_i^{tk}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}} \\
& = \frac{\sum_{i=1}^{N}\left(p_{ik}m_i^{sk}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{ik}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{is}m_i^{sk}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}} = PI\left(s,k;q_f^{sk}\right)^{\frac{\sigma-1}{\sigma}}
\end{aligned}
\tag{A.18}
$$

That is, we get

$$
\begin{aligned}
& \frac{\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}} \times \frac{\sum_{i=1}^{N}\left(p_{ik}m_i^{tk}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{ik}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{it}m_i^{tk}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}} \\
& = \frac{\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{it}m_i^{tk}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}} \times \frac{\sum_{i=1}^{N}\left(p_{ik}m_i^{tk}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{ik}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}} \\
& = \frac{\sum_{i=1}^{N}\left(p_{ik}m_i^{sk}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{ik}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{is}m_i^{sk}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}}.
\end{aligned}
\tag{A.19}
$$

Because the last line of the above equation does not depend on the prices in state t, the second line does not depend on the price vectors in state t, either. In other words, the following terms are used:

$$
\frac{\sum_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{it}m_i^{tk}\left(q_i\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}},
$$

does not depend on $p_{it.}$ This is possible only when we get

$$
m_i^{st}\left(q_i\right) = m_i^{tk}\left(q_i\right) \text{ for all } i = 1, 2, .., N, \text{ and } s, t, k = 1, 2, ..., M.
$$

Thus, the mean function, m_i, must take an identical value across all possible combinations of states, s, t. Thus, we obtain the following equations.

$$
m_i^{st}\left(q_i\right) = m_i^{tk}\left(q_i\right) = m_i\left(q_i\right).
$$

Assume $\sigma = 1$, then, from the sate reversal test, we must have

$$
\begin{aligned}
& PI\left(s,t;q_f^{ts}\right) \times PI\left(t,s;q_f^{ts}\right) \\
= & \frac{\prod_{i=1}^{N}\left(w_{it}\right)^{-w_{it}} \times \prod_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_i\right)\right)^{w_{it}}}{\prod_{i=1}^{N}\left(w_{is}\right)^{-w_{is}} \times \prod_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_i\right)\right)^{w_{is}}} \times \frac{\prod_{i=1}^{N}\left(w_{is}\right)^{-w_{is}} \times \prod_{i=1}^{N}\left(p_{is}m_i^{ts}\left(q_i\right)\right)^{w_{is}}}{\prod_{i=1}^{N}\left(w_{it}\right)^{-w_{is}} \times \prod_{i=1}^{N}\left(p_{it}m_i^{ts}\left(q_i\right)\right)^{w_{it}}} \\
= & 1
\end{aligned}
$$

Note that we can erase the price and share vectors, which leads to

$$
\frac{\prod_{i=1}^{N}\left(m_i^{st}\left(q_i\right)\right)^{w_{it}}}{\prod_{i=1}^{N}\left(m_i^{st}\left(q_i\right)\right)^{w_{is}}} \times \frac{\prod_{i=1}^{N}\left(m_i^{ts}\left(q_i\right)\right)^{w_{is}}}{\prod_{i=1}^{N}\left(m_i^{ts}\left(q_i\right)\right)^{w_{it}}} = 1,
$$

Using $m_i^{ts}\left(q_i\right) = m_i^{st}\left(q_{i\_st}\right),$

$$
\prod_{i=1}^{N}\left(\frac{m_i^{st}\left(q_i\right)}{m_i^{st}\left(q_{i\_st}\right)}\right)^{w_{it}} = \prod_{i=1}^{N}\left(\frac{m_i^{st}\left(q_i\right)}{m_i^{st}\left(q_{i\_st}\right)}\right)^{w_{is}}. \tag{A.20}
$$

Assume for some i, we have

$$
m_i^{st}\left(q_i\right) \neq m_i^{st}\left(q_{i\_st}\right). \tag{A.21}
$$

Then, by choosing w_{it} and w_{is} appropriately, it is possible to have

$$
\prod_{i=1}^{N}\left(\frac{m_i^{st}\left(q_i\right)}{m_i^{st}\left(q_{i\_st}\right)}\right)^{1/N} \neq \prod_{i=1}^{N}\left(\frac{m_i^{st}\left(q_i\right)}{m_i^{st}\left(q_{i\_st}\right)}\right)^{w_{is}}, \tag{A.22}
$$

which is a contradiction. Therefore, for all $i = 1, 2, .., N$, and $s, t = 1, 2, ..., M$,we must obtain

$$
m_i^{st}\left(q_i\right) = m_i^{st}\left(q_{i\_st}\right). \tag{A.23}
$$

By transitivity, we have

$$
\begin{aligned}
&PI\left(s,t;q_f\right)\times PI\left(t,k;q_f^{tk}\right)\\
=\ &\frac{\prod_{i=1}^{N}\left(w_{it}\right)^{-w_{it}}\times\prod_{i=1}^{N}\left(p_{it}m_i^{st}\left(q_i\right)\right)^{w_{it}}}{\prod_{i=1}^{N}\left(w_{is}\right)^{-w_{is}}\times\prod_{i=1}^{N}\left(p_{is}m_i^{st}\left(q_i\right)\right)^{w_{is}}}\times\frac{\prod_{i=1}^{N}\left(w_{ik}\right)^{-w_{ik}}\times\prod_{i=1}^{N}\left(p_{ik}m_i^{tk}\left(q_i\right)\right)^{w_{ik}}}{\prod_{i=1}^{N}\left(w_{it}\right)^{-w_{is}}\times\prod_{i=1}^{N}\left(p_{it}m_i^{tk}\left(q_i\right)\right)^{w_{it}}}\\
=\ &\frac{\prod_{i=1}^{N}\left(w_{ik}\right)^{-w_{sk}}\times\prod_{i=1}^{N}\left(p_{ik}m_i^{sk}\left(q_i\right)\right)^{w_{ik}}}{\prod_{i=1}^{N}\left(w_{is}\right)^{-w_{sk}}\times\prod_{i=1}^{N}\left(p_{is}m_i^{sk}\left(q_i\right)\right)^{w_{is}}}.
\end{aligned}
\tag{A.24}
$$

We can erase the price and share variables, which leads to the following simple relation,

$$
\frac{\prod_{i=1}^{N}\left(m_i^{st}\left(q_i\right)\right)^{w_{it}}}{\prod_{i=1}^{N}\left(m_i^{st}\left(q_i\right)\right)^{w_{is}}}\times\frac{\prod_{i=1}^{N}\left(m_i^{tk}\left(q_i\right)\right)^{w_{ik}}}{\prod_{i=1}^{N}\left(m_i^{tk}\left(q_i\right)\right)^{w_{it}}}=\frac{\prod_{i=1}^{N}\left(m_i^{sk}\left(q_i\right)\right)^{w_{ik}}}{\prod_{i=1}^{N}\left(m_i^{sk}\left(q_i\right)\right)^{w_{is}}}. \tag{A.25}
$$

Because the R.H.S. of the above equation does not depend on w_{it}, the L.H.S. does not depend on w_{it}, either. Thus, we must get

$$
\frac{m_i^{st}\left(q_i\right)}{m_i^{tk}\left(q_i\right)}=A>0 \tag{A.26}
$$

Therefore,

$$
m_i^{st}\left(q_i\right)=Am_i^{tk}\left(q_i\right) \tag{A.27}
$$

Set $q_i=1$. Then, we get

$$
m_i^{st}\left(1\right)=Am_i^{tk}\left(1\right)=1, \tag{A.28}
$$

$$
A=1 \tag{A.29}
$$

Therefore, we obtain

$$
m_i^{st}\left(q_i\right)=m_i^{tk}\left(q_i\right)=m_i\left(q_i\right) \tag{A.30}
$$

■

A.3 Proof for Proposition 9

Proof. The price index between the two states, $s, t,$ is,

$$PI\left(p_s, p_t, q\right) = \frac{\left(\sum_{i=1}^{N}\left(p_{it}m_i\left(q_{i1}, q_{i2}, ..., q_{iM}\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}}{\left(\sum_{i=1}^{N}\left(p_{is}m_i\left(q_{i1}, q_{i2}, ..., q_{iM}\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}} \text{ if } \sigma \neq 1 \tag{A.31}$$

$$= \frac{\prod_{i=1}^{N}\left(w_{it}\right)^{-w_{it}} \times \prod_{i=1}^{N}\left(p_{it}m_i\left(q_{i1}, q_{i2}, ..., q_{iM}\right)\right)^{w_{it}}}{\prod_{i=1}^{N}\left(w_{is}\right)^{-w_{is}} \times \prod_{i=1}^{N}\left(p_{is}\left(q_{i1}, q_{i2}, ..., q_{iM}\right)\right)^{w_{is}}} \text{ if } \sigma = 1. \tag{A.32}$$

Suppose $\sigma = 1$. Let us change the quantities in state 1 by λ. Then, by the invariance to proportional changes in a state, we obtain

$$\frac{\prod_{i=1}^{N}\left(p_{it}m_i\left(q_{i1}, q_{i2}, ..., q_{iM}\right)\right)^{w_{it}}}{\prod_{i=1}^{N}\left(p_{is}\left(q_{i1}, q_{i2}, ..., q_{iM}\right)\right)^{w_{is}}} = \frac{\prod_{i=1}^{N}\left(p_{it}m_i\left(\lambda q_{i1}, ..\right)\right)^{w_{it}}}{\prod_{i=1}^{N}\left(p_{isi}m_i\left(\lambda q_{i1}, ..\right)\right)^{w_{is}}} \tag{A.33}$$

We can simplify the above to

$$\frac{\prod_{i=1}^{N}\left(m_i\left(q_{i1}, ...\right)\right)^{w_{it}}}{\prod_{i=1}^{N}\left(m_i\left(q_{i1}, ...\right)\right)^{w_{is}}} = \frac{\prod_{i=1}^{N}\left(m_i\left(\lambda q_{i1}, ..\right)\right)^{w_{it}}}{\prod_{i=1}^{N}\left(m_i\left(\lambda q_{i1}, ..\right)\right)^{w_{is}}}. \tag{A.34}$$

Arranging the above,

$$\frac{\prod_{i=1}^{N}\left(m_i\left(\lambda q_{i1}, ..\right)\right)^{w_{it}}}{\prod_{i=1}^{N}\left(m_i\left(q_{i1}, ...\right)\right)^{w_{it}}} = \frac{\prod_{i=1}^{N}\left(m_i\left(\lambda q_{i1}, ..\right)\right)^{w_{is}}}{\prod_{i=1}^{N}\left(m_i\left(q_{i1}, ...\right)\right)^{w_{is}}}. \tag{A.35}$$

Because the R.H.S does not depend on w_{it},while the L.H.S. does not depend on w_{is}, $(A.35)$ does not depend on w_{it}. Therefore, we should be able to find a function L such that

$$\frac{\prod_{i=1}^{N}\left(m_i\left(\lambda q_{i1},..\right)\right)^{w_{it}}}{\prod_{i=1}^{N}\left(m_i\left(q_{i1},...\right)\right)^{w_{it}}}=\frac{L_t\left(\lambda q_{i1},..\right)}{L_t\left(q_{i1},...\right)}, \tag{A.36}$$

$$\frac{\prod_{i=1}^{N}\left(m_i\left(\lambda q_{i1},..\right)\right)^{w_{is}}}{\prod_{i=1}^{N}\left(m_i\left(q_{i1},...\right)\right)^{w_{is}}}=\frac{L_s\left(\lambda q_{i1},..\right)}{L_s\left(q_{i1},...\right)}. \tag{A.37}$$

Thus, we can apply Lemma 7 to get a function $f\left(\lambda\right)$ such that

$$\frac{\prod_{i=1}^{N}\left(m_i\left(\lambda q_{i1},..\right)\right)^{w_{it}}}{\prod_{i=1}^{N}\left(m_i\left(q_{i1},...\right)\right)^{w_{it}}}=\frac{\prod_{i=1}^{N}\left(m_i\left(\lambda q_{i1},..\right)\right)^{w_{is}}}{\prod_{i=1}^{N}\left(m_i\left(q_{i1},...\right)\right)^{w_{is}}}=f\left(\lambda\right) \tag{A.38}$$

Take the natural logarithms,

$$\begin{aligned}\sum_{i=1}^{N} w_{it}\left(\ln m_i\left(\lambda q_{i1},..\right)-\ln m_i\left(q_{i1},...\right)\right) &= \sum_{i=1}^{N} w_{is}\left(\ln m_i\left(\lambda q_{i1},..\right)-\ln m_i\left(q_{i1},...\right)\right)\\ &= \ln f\left(\lambda\right).\end{aligned} \tag{A.39}$$

Using the fact that

$$\sum_{i=1}^{N} w_{it}=1,$$

we can obtain

$$\begin{aligned}\sum_{i=1}^{N} w_{it}\left(\ln m_i\left(\lambda q_{i1},..\right)-\ln m_i\left(q_{i1},...\right)-\ln f\left(\lambda\right)\right) &= \sum_{i=1}^{N} w_{is}\left(\ln m_i\left(\lambda q_{i1},..\right)-\ln m_i\left(q_{i1},...\right)-\ln f\left(\lambda\right)\right)\\ &= 0\end{aligned} \tag{A.40}$$

Because the above is an identity with respect to w_{it} an w_{is}, we get

$$\ln m_i\left(\lambda q_{i1},..\right)-\ln m_i\left(q_{i1},...\right)-\ln f\left(\lambda\right)=0.$$

That is,

$$m_i\left(\lambda q_{i1},..\right)=f\left(\lambda\right)m_i\left(q_{i1},...\right) \tag{A.41}$$

Next, suppose $\sigma \neq 1$. Again, let us change the quantities in state 1 by λ. Then, the invariant property implies that:

$$\frac{\sum_{i=1}^{N}\left(p_{it}m_i\left(q_{i1},...\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{is}m_i\left(q_{i1},...\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}}=\frac{\sum_{i=1}^{N}\left(p_{it}m_i\left(\lambda q_{i1},..\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{is}m_i\left(\lambda q_{i1},..\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}}. \tag{A.42}$$

Arranging the above,

$$\frac{\sum_{i=1}^{N}\left(p_{it}m_i\left(\lambda q_{i1},..\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{it}m_i\left(q_{i1,},...\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{it}\right)^{\frac{1}{\sigma}}}=\frac{\sum_{i=1}^{N}\left(p_{is}m_i\left(\lambda q_{i1},..\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{is}m_i\left(q_{i1},...\right)\right)^{\frac{\sigma-1}{\sigma}}\left(w_{is}\right)^{\frac{1}{\sigma}}}. \tag{A.43}$$

Because the R.H.S. does not depend on p_{it}, while the L.H.S. does not depend on p_{is}, there must exist a function, $g\left(\lambda,\sigma\right)$ such that

$$\frac{\sum_{i=1}^{N}\left(p_{it}m_i\left(\lambda q_{i1},..\right)\right)^{\frac{\sigma-1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{it}m_i\left(q_{i1},...\right)\right)^{\frac{\sigma-1}{\sigma}}}=\frac{\sum_{i=1}^{N}\left(p_{is}m_i\left(\lambda q_{i1},..\right)\right)^{\frac{\sigma-1}{\sigma}}}{\sum_{i=1}^{N}\left(p_{is}m_i\left(q_{i1},...\right)\right)^{\frac{\sigma-1}{\sigma}}}=g\left(\lambda,\sigma\right), \tag{A.44}$$

which leads to the following equation,

$$\sum_{i=1}^{N}\left(p_{it}m_i\left(\lambda q_{i1},..\right)\right)^{\frac{\sigma-1}{\sigma}}=g\left(\lambda,\sigma\right)\sum_{i=1}^{N}\left(p_{it}m_i\left(q_{i1},...\right)\right)^{\frac{\sigma-1}{\sigma}}. \tag{A.45}$$

Arranging the above,

$$\sum_{i=1}^{N}p_{it}^{\frac{\sigma-1}{\sigma}}\left(\left(m_i\left(\lambda q_{i1},..\right)\right)^{\frac{\sigma-1}{\sigma}}-g\left(\lambda,\sigma\right)\left(m_i\left(q_{i1},...\right)\right)^{\frac{\sigma-1}{\sigma}}\right)=0. \tag{A.46}$$

Because the above is an identity with respect to p_{it}, we must obtain

$$\left(m_i\left(\lambda q_{i1},..\right)\right)^{\frac{\sigma-1}{\sigma}}=g\left(\lambda,\sigma\right)\left(m_i\left(q_{i1},...\right)\right)^{\frac{\sigma-1}{\sigma}}. \tag{A.47}$$

Take the power to the $\frac{\sigma}{\sigma-1}$, and arrange the above function,

$$\frac{m_i\left(\lambda q_{i1},..\right)}{m_i\left(q_{i1},...\right)}=g\left(\lambda,\sigma\right)^{\frac{\sigma}{\sigma-1}}. \tag{A.48}$$

Because the L.H.S. does not depend on $\sigma, g\left(\lambda,\sigma\right)^{\frac{\sigma}{\sigma-1}}$ does not depend on σ, either. Therefore, we define

$$f\left(\lambda\right)=g\left(\lambda,\sigma\right)^{\frac{\sigma}{\sigma-1}}. \tag{A.49}$$

Then, we get

$$m_i\left(\lambda q_{i1},..\right)=f\left(\lambda\right)m_i\left(q_{i1}...\right), \tag{A.50}$$

which is identical to $(A.41)$.

From Theorem 1 in Luce (1964), p.281, if the functions m_i and f are increasing functions, the general solution of the functional equation is given by

$$m_i\left(q_{i1}, q_{i2}, q_{i3}, , .., q_{iM}\right) = \prod_{m=1}^{M} \left(q_{im}\right)^{c_i}. \tag{A.51}$$

From Proposition 5, we get

$$c_i = 1/M.$$

Therefore, for all i, we obtain

$$m_i\left(q_{i1}, q_{i2}, q_{i3}, , .., q_{iM}\right) = \prod_{m=1}^{M} \left(q_{im}\right)^{1/M}. \tag{A.52}$$

■

B Data Appendix

B.1 Weekly Scanner Data

We use weekly Japanese store-level scanner data, known as SRI+, collected by INTAGE Inc. The data set includes sales records of processed foods, daily necessities, cosmetics, and drugs with a Japanese Article Number (JAN) code. The data set covers about 3,000 stores, located all over Japan.

The price indexes and cost-of-living indexes were computed using data for the category of ice cream sold in supermarkets for the week of January 8, 2017 through the week of September 11, 2022. The stores from which data were collected were limited to the 960 stores where ice cream sales were recorded for all weeks of the relevant period. The number of ice cream brands recorded for the relevant period is 4843. When the same brand is sold in different stores, the number of items is 827,919 if they are considered to be different products.

B.2 Monthly Official Statistics

In Section 7.2, two publicly available official statistics, (1) the Retail Price Survey (hereafter, RPS), and (2) the Family Income and Expenditure Survey (hereafter, FIES), were used to estimate the price index for oranges. We begin with an overview of these official statistics, followed by a description of how the data set for oranges was created.

The Consumer Price Index in Japan is estimated by collecting price information from RPS and expenditure information from FIES used as weights . The former is a survey of stores and the latter is a survey of households; both are monthly surveys. As shown below, the two surveys are closely related. First, both surveys are conducted by the Statistics Bureau of the Ministry of Internal Affairs and Communications (hereafter, Statistics Bureau of Japan), and the target municipalities were selected using the same sampling method. Specifically, of the 1,718 municipalities in Japan, 168 are selected by both statistics. Second, the selection of survey items in RPS is based on information from FIES. In other words, whether an item reaches at least 1/10,000th of the total expenditure in FIES is a criterion for its selection as a survey item in RPS. A summary of each survey is as follows.

1) **Retail Price Survey (RPS):**

   How to conduct the survey

   The investigator goes to the target store and asks for the price of the product designated as the basic brand. Price information is entered into a tablet terminal by the investigator and transmitted to the Statistics Bureau of Japan.

   How to select stores for survey

   The number of stores surveyed is approximately 28,000 throughout Japan. The survey stores are representative stores with high sales volume in the survey area, and the selection method is not random.

   Period of the survey

   The survey is conducted on one Wednesday, Thursday or Friday of the week that includes the 12th of each month. However, some fresh products are surveyed a total of three times: early (week including the 5th), mid (week including the 12th), and late (week including the 22nd).

   Items covered in the survey

   The number of items surveyed is approximately 500, although it varies from month to month. The prices to be published are those for commodities designated as basic brand, with no quality adjustments implemented. Note that unit prices are published for fresh products.

   Data availability

   Monthly and item-specific price information is published for 81 cities. Incidentally, quality-adjusted price indexes by item are published only for the nation and for the Tokyo metropolitan area. Therefore, in order to obtain quality-adjusted price information on a prefectural basis, econometrician must conduct their own quality adjustment from RPS, which is a very difficult task.

2) **Family Income and Expenditure Survey (FIES):**

   How to conduct the survey

   Households respond to the following questions: (a) demographics of household members; (b) income for the past year; and (c) diary-based household income and expenditures. The diaries here are specially prepared by Statistics Bureau of Japan, and surveyors visit the surveyed households to provide detailed instructions on how to fill out the form. With these survey methods, it is assumed that the information on expenditures obtained from FIES is considerably more accurate than the recall method.

   How to select households for survey

Surveyed households are selected using a stratified three-stage sampling method (municipality - unit district - household). As mentioned above, the municipalities selected in the first stage are the same as those covered by RPS.

The total number of households with two or more persons is approximately 8,000.

Period of the survey

As a rule, a household is surveyed continuously for six months. However, the survey period for single-person households is three months. Every month, one-sixth of the households are replaced.

Items covered in the survey

There are no pre-defined items. This is due to the fact that the item column of the survey form is an open-ended field, as households spend on a wide variety of goods and services. In addition, by not pre-setting the survey items, it will be possible to collect information on new products.

Data availability

Although there are no pre-defined items, the Statistics Bureau of Japan tabulates entries in diary into broad categories or item units, and publishes information on expenditures by approximately 500 items. Information on expenditures by month and by item is published only for households of two or more persons in 50 cities, including prefectural capitals.

3) **How to merge RPS and FIES:**

In this paper, we compiled monthly panel data by 47 prefectural capitals and by item from the above two surveys. However, the two surveys have different item codes, and furthermore, the item contents are not exactly the same. Therefore, when the researcher independently matches the two surveys, it is necessary to split or combine the items. The method of dividing and combining items is partially published by Statistics Bureau of Japan at the following website (however, the PDF file is written in Japanese).

Statistics Bureau of Japan (2022) "Division and integration of the FIES items to the CPI items," 2020-Base Explanation of the Consumer Price Index, IV 6.

URL: https://www.stat.go.jp/data/cpi/2020/kaisetsu/pdf/4-6.pdf

4) **Reasons for choosing the item "oranges"**

We chose fresh fruit as our estimation target because of its seasonality and lack of need for quality adjustment. Among fresh fruits, we reported price indexes for oranges that meet the following conditions. The first condition is that the price data exist in January 2018, the base date. The second condition is that price information is available in at least one of the 47 prefectures for each month. Also, because the two surveys are combined, there are cases where price information exists even when there is zero spending, or where there is positive spending but the price information is missing. In this paper, data with zero expenditures or missing price information were removed.

5) **Estimated results for σ:**

When estimating the elasticity of substitution (σ) based on the monthly panel data by item for 47 prefectures, we use Feenstra's (1994) method. The results are shown in the following appendix table. Note that the four items listed in the table are those that satisfy the two conditions described in (4) above.

Table A1: Estimation Results for σ

Item	Estimated σ
Oranges	9.06
Apples	8.87
Bananas	2.30
Kiwi fruits	6.94

B.3 Data from the International Comparison Program (ICP) 2017 for Cross-country Comparisons of Price Levels and Real Expenditures

We use data from the ICP covering 177 countries of the world. Details of the methods used in the compilation of data and the subsequent compilation of PPPs and real expenditures at the regional and global level can be found in the 2017 ICP reports from the World Bank (2020). In this paper, we focus on data for household consumption which consists of 109 basic headings.